EUROPEAN ORGANISATION
FOR THE SAFETY OF AIR NAVIGATION

EUROCONTROL

EUROCONTROL EXPERIMENTAL CENTRE

MONOTONY IN AIR TRAFFIC CONTROL - CONTRIBUTING FACTORS AND MITIGATION STRATEGIES

EEC Note No. 15/06

Project SAS-2-HF-AAAA

Issued: November 2006

REPORT DOCUMENTATION PAGE

Reference: EEC Note No. 15/06	**Security Classification:** Unclassified
Originator: EEC – CoE **SAS** (**S**afety **A**nalytic and **S**cientific)	**Originator (Corporate Author) Name/Location:** EUROCONTROL Experimental Centre Centre de Bois des Bordes B.P.15 F - 91222 Brétigny-sur-Orge Cedex FRANCE Telephone : +33 (0)1 69 88 75 00Internet : www.eurocontrol.int
Sponsor: EUROCONTROL Human Factors Laboratory	**Sponsor (Contract Authority) Name/Location:** EUROCONTROL Experimental Centre Centre de Bois des Bordes B.P.15 F - 91222 Brétigny-sur-Orge Cedex FRANCE Telephone: +32 2 729 9011 www.eurocontrol.int

TITLE:

MONOTONY IN AIR TRAFFIC CONTROL - CONTRIBUTING FACTORS AND MITIGATION STRATEGIES

Authors	Date	Pages	Figures	Tables	Annexes	References
Sonja Straussberger	11/2006	xlvi + 272	66	155	4	467

EEC Contact	Project	Task No. Sponsor	Period
Dirk Schaefer	SAS-2-HF-AAAA		2003 to 2005

Distribution Statement:

(a) Controlled by: Head of SAS
(b) Special Limitations: None

Descriptors (keywords):

ATC - Monotony - Repetitiveness - Critical States - Boredom - Underload - Stress -Traffic Load - Dynamic Density - Traffic Synchronization - Psychophysiology - Field Study – Simulation.

Abstract:
This report discusses the role of monotony in Air Traffic Control (ATC). Despite its obvious relevance as a critical individual state in air traffic controllers, monotony has not been well researched in the past. To describe evoking and contributing factors, distinguish similar critical states such as fatigue and satiation, and define countermeasures, three experimental studies were conducted with a total of 32 air traffic controllers (ATCO) in the simulated and ten ATCOs in an operational air traffic control environment. Traffic repetitiveness and (dynamic) traffic density were confirmed to evoke a state of monotony, which is indicated in reduced physiological activation, subjective sleepiness, and behavioral impairments. At the same time, reduced workload but also impaired cognitive functions were observed while fatigue increased with higher time-on-task. Higher initial recovery, the experience of flow, (dynamic) traffic density changing from low to high and active physical exercises in rest breaks were determined to have a monotony-reducing effect. Based on these outcomes, recommendations address the assessment procedures during ATC concept development as well as options for the improvement of the operational environment. The applied psychophysiological multilevel-assessment method shows otherwise undetected but critical dissociations as related to the experience of cognitive functions and motivational aspects and suggests the application of assessment procedures beyond workload ratings. In the operational environment, systematic position assignment based on predicted traffic changes, the collection of initial state information, balanced active rest breaks, trainings on the role of mental sets, and the consideration of an ATCOs psychophysiological condition in incident reporting systems are proposed. A model integrating the mentioned factors supports a systematic analysis of this issue. Future research may address the role of further individual factors related to personnel selection and the long-term development of critical states.

FOREWORD

There are various problems associated with monotony in air traffic management.

The first and most obvious problem is that monotony will continue to pose a challenge to the work of air traffic controllers and thus a threat to aviation safety. One could argue that increases in air traffic demand and the associated increase in controller workload will render this problem obsolete yet, unfortunately, this is not the case. Monotony due to very low traffic load will continue to play a role for example in night shifts, resulting in problems of vigilance. More importantly perhaps, monotony related to repetitive traffic patterns may play an increasing role since the need to manage higher traffic loads in busy periods might lead to increasingly uniform traffic patterns. Future levels of automation could further aggravate this problem.

The second and perhaps less obvious problem is that to date we have a limited understanding of what monotony actually is, particularly in the area of air traffic management. The construct itself is not very well defined: the term monotony is sometimes used as referring to an operator's state and sometimes to denote a situation inducing such a state. The factors contributing to an operator state of monotony are not very well understood either. And finally the consequences of monotony in terms of air traffic controllers' performance are not fully understood.

A further problem might become more apparent as soon as the factors contributing to monotony and the impact of monotony on operator performance and well-being will be better researched, namely the prevention and mitigation of monotony. The first solution that comes to mind would be to avoid work situations proven to increase monotony and impair performance. However, in some instances that may be either impractical or simply impossible. Whilst there is a sense that mitigation may help in such cases we have a very poor understanding of how we can mitigate monotony.

For the above reasons EUROCONTROL has decided to sponsor research in the field of monotony through a Ph.D. scholarship for Sonja Straussberger. I had the privilege and pleasure to supervise Sonja's Ph.D. thesis at EUROCONTROL. I am very pleased with the results of her research and I am convinced they will make a significant contribution to ATM research.

Dirk Schaefer
EEC Quality Manager

Page intentionally left blank

ACKNOWLEDGEMENTS

In a marvelous description of the history of experimental psychology in an edition published in 1956, Edward Boring determined two limitations for scientific advance. The first one is ignorance, which means that discovery depends on another discovery being made to open its way. The second one is when discovery is limited by the habits of thought that pertain to the culture of any region or period. Boring called this phenomenon *Zeitgeist*.

With this statement in mind, I am in the pleasant situation to acknowledge the aid and cooperation of numerous persons who enabled research on a subject which was trying to overcome these limitations.

First of all, I am indebted to my supervisors, Professor K. Wolfgang Kallus, Department of Work-, Organizational and Environmental Psychology at the Karl-Franzens-University Graz, for the academic supervision, and Dr. Dirk Schaefer, EUROCONTROL Experimental Centre, Brétigny sur Orge, for the industrial supervision. I was fortunate to have the inspiring guidance and encouragement of two exceptional thinkers, who made it possible to arrange this research thanks to their unbureaucratic, flexible, and open attitudes.

Furthermore I owe a very special vote of thanks to Professor Dr. Peter Jorna and Mr. Alistair Jackson for taking their time for encouraging statements, the reading and assessing processes and the discussion. Their creative minds gave me valuable insights in relevant relationships.

My expression of thanks is long over-due to numerous air traffic controllers and experts who patiently and curiously participated in the data collection, the study preparation or the completion process. This also includes the persons who enabled access to control centers or volunteered in the local organization of the studies. Also, without the kind support of many colleagues at EUROCONTROL the current work would not have been possible.

This research was funded by EUROCONTROL Experimental Centre and enabled by the progressive engagement of proactive people who made this work possible. Hopefully, many students will be able to make use of such an opportunity in future.

Finally, special thanks are due to those who supported me personally. My partner Laurent for his never-ending understanding, my family for supporting my career, and my friends for who distance does not matter and who have been there to help out any time.

The opportunity to conduct the studies autonomously and learn greatly contributed to develop a better understanding of science in an applied field and to take the responsibility work psychologists have representing the human component in a technical field. More than anything, the work reinforced my understanding of the significance of scientific ethics in applied environments. This approach was also significantly influenced by the critical discussion of Schmidtke in 2002 of research trends in the field of ergonomics and the present disinformation in publications. He emphasized the responsibility of the researcher to carefully apply scientific methods for the scope of increasing knowledge that can be applied to solve practical problems.

Certainly, a lot of people are involved in the completion of such a big work. As it is not possible to mention all of them by name on one page, everyone else who is concerned definitely will know that these special thanks are TO YOU.

Page intentionally left blank

SUMMARY

Despite the obvious relevance of monotony in the field of air traffic control (ATC), it has not been well researched in the past. One of the reasons is related to the unclear use of monotony to designate the task characteristics as well as the individual reactions to these characteristics. At the same time, the more frequently addressed concepts of boredom, underload, and low vigilance were not clearly distinguished and kept apart from critical states such as fatigue and satiation. While these states are similar in appearance, the occurrence of either state requires different countermeasures. Moreover, existing research results cannot be directly applied to understand monotony in ATC, as they were obtained in different industries or focused on isolated components. In addition, previously developed models to explain the effects of task execution on performance and individual states are often restricted on few components and thus do not well represent the conditions contributing to monotony.

For the current work, a framework was used that clearly distinguished between the description of task characteristics in terms of uneventfulness and repetitiveness and exclusively used monotony for the description of an individual state. Based on research results obtained by Bartenwerfer (1957), monotony was characterized by physiological deactivation, increased feelings of tiredness and boredom, and fluctuating performance. Under consideration of individual and organizational factors, the framework allowed not only investigating the effect of task characteristics, but also the distinction of critical states through the assessment of multiple indicators on a physiological, subjective, and behavioral level.

To determine task and individual factors that evoke, enhance or mitigate monotony in ATC, to distinguish critical states and to define countermeasures, three studies were conducted. In simulated air traffic control settings a small-scale experiment with eight operational experts (not active controllers) and a main study with 24 air traffic controllers were run. Ten air traffic controllers participated in a field study executed in an European Control Center. As repetitiveness is an important component not only in the current air traffic management (ATM), but also in future concepts favoring air traffic synchronization, it was centered in the research activities. In addition, it was assumed that repetitiveness might have a different impact depending on the level of dynamic traffic density.

The main simulation experiment was based on the small-scale study and involved a 2 (break activity) x 2 (repetitiveness) x 2 (sequence of dynamic density) x 2 (run) x 3vs.15 (interval) -mixed design with repeated measures on the last two factors. Two traffic scenarios of 45 minutes each were executed and a short third scenario was introduced to determine the effects of break activity. The dependent variables comprised heart rate (HR) and its variability (HRV), skin conductance level, blink rate, and the power in common frequency bands of spontaneous brain activity. On a subjective level, scales assessed mood, workload, and the perceived cognitive, emotional and motivational state during and after the scenarios. Behavioral and performance measures assessed the occurrence of Short Term Conflict Alerts (STCA). HR, sleepiness, and the subjective feeling of monotony were integrated in a standardized indicator for the state of monotony based on the small-scale study. Higher monotony occurred if participants were exposed to repetitive scenarios. The effect of monotony was reinforced in the low density condition of the first run and also reflected in tendencially increased conflict resolution time in an unexpected situation. The comparison of indicators for the critical states revealed that monotony as a consequence of task repetitiveness was clearly found in the first scenario, but overlaid by time-on-task effects resulting in higher fatigue with the ongoing second scenario. The distinction of critical states did not allow a clear statement concerning satiation. While the sequence of dynamic density changing from high to low from the first to the second run still increased the cognitive impairments, a motivating and monotony-decreasing effect of the dynamic density changing from low to high was found.

The monotony-decreasing effect of active exercises in rest breaks was confirmed, even though there was no favoring effect after repetitive conditions. Boredom proneness and initial recovery and strain states were not confirmed to be a significant factor contributing to monotony; marginally significant effects do however indicate the relevance of further investigation. On the other hand, if individuals perceived flow during task execution in the first run, the indicator for the state of monotony was lower.

In the operational environment, a 2 x 2 within-subject design was deployed with high versus low repetitiveness and high versus low traffic load in sectors. Controllers participated in 90-minute-work periods that had been selected based on supervisors' ratings and traffic statistics. Physiological indicators comprised HR and HRV, the previously used subjective scales and questionnaires were extended with ratings for traffic characteristics, and performance indicators were collected through subjective ratings of related behaviors. The effects of repetitiveness on the composed indicator for the state of monotony were confirmed. A more detailed analysis revealed that – in contrary to the simulated environment - the effects were not reflected in the summarized physiological measures during a work period. Controllers experienced reduced motivation, attentiveness, concentration, and increased boredom, but also reduced workload and strain. Some of these effects were even more pronounced in the low traffic load condition. Apart from that, subjectively perceived motivation and the combined indicator for the state of monotony were higher if a change of traffic density from low to high was perceived during a work period. Nonetheless, the description of individual cases showed covered physiological effects which turned out to be rather the consequence of clearly distinguishable events on the individual level. Delayed and immediate effects on blood pressure were observed under consideration of personally relevant occurrences. On an individual level the initial state of recovery at the beginning of the work day was confirmed to influence the development of critical states. At the same time the collection of organizational processes helped to understand changes in subjectively perceived satiation.

The total of the results supports the assumption that repetitiveness in task conditions is evoking monotony in both simulation and field settings, which is mitigated by the state of recovery at the beginning of the work shift. The potential influence of boredom proneness and the unexpected effect of flow experience require further investigation. The contradictive results in the physiological indicators are explained by behaviors executed by air traffic controllers to remain active. The results do not support any interpretation related to stress, as - opposed to research studies that used ATC-related tasks - workload was also reduced. This led to propose a model of monotony that considers the task factors repetitiveness and uneventfulness, the individual boredom proneness and states at the beginning of the work shift as well as organizational factors to assess monotony with the help of physiological, subjective, and behavioral indicators. The distinction of other states such as fatigue and satiation and a positive state of flow is essential, even though with the current data the definition of satiation remains unclear.

Based on these outcomes, recommendations address the level of ATC concept development as well as the improvement of the operational environment. Several methodological issues are stressed to be considered in simulation set-ups. They contain the multi-level approach to assess controller states as a task consequence, the selection of increased scenario duration, and sufficient training in new concepts. Especially a one-sided assessment of workload ignores further negative effects as related to cognitive functioning and motivational aspects. In the operational environment, the systematic consideration of changes in traffic density and collection of initial state information call for a systematic assignment of controllers to work positions based on traffic predictions to make use of their motivating and monotony counteracting potential. Trainings may include further sensitization towards the effect of habits and mental sets and also provide better strategies for balanced rest breaks and systematic communication. Finally, the collected statistics in incident reporting systems should be extended by better categorized information on the controller's psychophysiological states. Overall, future questions may address the role of further individual factors related to personnel selection and the long-term development of critical states. Hence, monotony remains a challenging issue within the ATC environment.

TABLE OF CONTENTS

LIST OF APPENDICES

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS AND ACRONYMS

ANOVA	**An**alysis **o**f **Va**riance
ANS	**A**utonomic **N**ervous **S**ystem
ANSP	**A**ir **N**avigation **S**ervice **P**rovider
APP	**App**roach Control
ARAS	**A**scending **R**eticular **A**ctivating **S**ystem
ATC	**A**ir **T**raffic **C**ontrol
ATCO	**A**ir **T**raffic **Co**ntroller
ATM	**A**ir **T**raffic **M**anagement
Bpm	**B**eats **p**er **M**inute
CFMU	**C**entral **F**low **M**anagement **U**nit
CNS	**C**entral **N**ervous **S**ystem
CWP	**C**ontroller **W**orking **P**osition
DBP	**D**iastolic **B**lood **P**ressure
DD	**D**ynamic **D**ensity
DDA	**D**escriptive **D**ata **A**nalysis
df	**D**egrees of **F**reedom
EC	**E**xecutive **C**ontroller
EDA	**E**lectrodermal **A**ctivity
eDEP	**E**arly **D**emonstration and **E**valuation **P**latform
EEC	**E**UROCONTROL **E**xperimental **C**entre
EEG	**E**lectrocardiogram
EOG	**E**lectrooculogram
ERP	**E**vent-**r**elated **P**otentials
HPA	**H**ypothalamus-**P**ituitary-**A**drenal-System
HR	**H**eart **R**ate
HRV	**H**eart **R**ate **V**ariability
IBI	**I**nter-**B**eat-**I**nterval
ICAO	**I**nternational **C**ivil **A**viation **O**rganization
ISO	**I**nternational **O**rganization for **S**tandardization
M	**M**ean Value
MANOVA	**M**ultivariate **A**nalysis of **V**ariance
MART	**Ma**lleable **R**esource **T**heory
MWL	**M**ental **w**ork**l**oad
OS	**O**ver**s**timulation
PAT	**P**eripheral **A**rterial **T**one
PC	**P**lanning **C**ontroller
RVSM	**R**educed **V**ertical **S**eparation **M**inima
SA	**S**ituation **A**wareness
SAM	**S**ympathetic-**A**dreno-**M**edullary-System
SBP	**S**ystolic **B**lood **P**ressure

SCL	**S**kin **C**onductance **L**evel
SD	**S**tandard **D**eviation
SDANN	**S**tandard **D**eviation of **n**ormal-to-**n**ormal Beats in 5-minute-intervals
SDNN	**S**tandard **D**eviation of **n**ormal-to-**n**ormal Beats
SDT	**S**ignal **D**etection **T**heory
STCA	**S**hort **T**erm **C**onflict **A**lert
US	**U**nder**s**timulation
VDT	**V**isual **D**isplay **T**ask
VTS	**V**ienna **T**est **S**ystem

REFERENCES

1. Abt, K. (1987). Descriptive Data Analysis: A concept between confirmatory and exploratory data analysis. *Methods of Information in Medicine*, 26, 77-88.

2. Åhsberg, E. (2000). Dimensions of fatigue in different working populations. *Scandinavian Journal of Psychology*, 41, 231–241.

3. Andreassi, J. L. (2000). *Psychophysiology. Human Behavior & Physiological Response* (3rd ed.). Hillsdale, New Jersey: Erlbaum.

4. Annett, J. (2002). Subjective rating scales: science or art? *Ergonomics*, 45(14), 966-987.

5. App, E. & Debus, G. (1998). Saccadic velocity and activation: development of a diagnostic tool for assessing energy regulation. *Ergonomics*, 41(5), 689-697.

6. Athènes, S., Averty, P., Puechmorel, S., Delahaye, D., & Collet, C. (2002). *ATC complexity and Controller Workload: Trying to bridge the gap*. HCI-Aero 2002, Boston (USA).

7. Averty, P., Athènes, S., & Collet C. (2002). *Evaluating a new index of mental workload in real control situation using psychophysiological measures*. 21st Digital Avionics Systems Conference proceedings DASC2002 (7a4p1-13), Irvine, USA, October 2002.

8. Backs, R. W. (1998). A comparison of factor analytic methods of obtaining cardiovascular autonomic components for the assessment of mental workload. *Ergonomics*, 41(5), 733-745.

9. Backs, R. W. (2001). An autonomic space approach to the psychophysiological assessment of mental workload. In P.A. Hancock & P. Desmond (Eds.). *Stress, workload and fatigue* (279-289). Mahwah: Erlbaum.

10. Backs, R. W., Lenneman, J. K., & Sicard, J. L. (1999). The Use of Autonomic Components to Improve Cardiovascular Assessment of Mental Workload in Flight Simulation. *The International Journal of Aviation Psychology*, 9(1), 33-47.

11. Backs, R. W., Navidzadeh, H. T. & Xu, X. (2000). Cardiorespiratory indices of mental workload during simulated air traffic control. *Proceedings of the IEA 2000/HFES 2000 Congress Volume 3*, Santa Monica, CA: Human Factors and Ergonomics Society, 89-92.

12. Badham, R., Clegg, C.W., & Wall, T.D. (2000). Sociotechnical theory. In W. Karwowski. (Ed.), *Handbook of Ergonomics.* New York: Wiley.

13. Bailey, J.P., Thackray, R.I., Pearl, J. & Parish, T.S. (1976). Boredom and arousal. Comparison of tasks differing in visual complexity. *Perceptual and Motor Skills*, 43, 141-142.

14. Baldamus, W. (1951). Type of work and motivation. *British Journal of Sociology*, 2, 44-58.

15. Baldamus, W. (1961). Efficiency and Effort. London: Tavistock. In Davies, 1961.

16. Banbury, S. & Tremblay, S. (2004). *A cognitive approach to situation awareness: theory and application. Aldershot*, UK: Ashgate.

17. Barmack, J. E. (1939a). Studies on the psychophysiology of boredom: Part I. The effect of 15 Mgs. of benzedrine sulfate and 60 Mgs. of ephedrine hydrochloride on blood pressure, report of boredom, and other factors. *Journal of Experimental Psychology*, 25, 494-505.

18. Barmack, J. E. (1939b). Studies on the psychophysiology of boredom: Part II. The effect of lowered room temperature and an added incentive on blood pressure, report of boredom and other factors. *Journal of Experimental Psychology*, 25, 634-642.

19. Barmack, J. E. (1939c). A definition of boredom: A reply to Mr. Berman. *American Journal of Psychology*, 52, 467-471.

20. Barmack, J. E. (1939d).The length of the work period and the work curve. *Journal of Experimental Psychology*, 25, 109-115.

21. Bartenwerfer, H. (1957). *Ueber die Auswirkungen einfoermiger Arbeitsvorgaenge* [The effects of uniform working conditions]. Untersuchungen zum Monotonieproblem vorgelegt von Heinrich Dueker in der Sitzung vom 25. Februar 1958. Marburg: Elwert.

22. Bartenwerfer, H. (1960). Untersuchungen zum Monotonieproblem [Experiments on the problem of monotony]. *Zentralblatt fuer Arbeitswissenschaft*, 14, 29-33.

23. Bartenwerfer, H. (1961). Psychische Beanspruchung [Mental Workload]. In A. Mayer & B. Herwig (Eds.*). Handbuch der Psychologie*, Betriebspsychologie (Vol 9, pp. 252-275). Goettingen: Hogrefe.

24. Bartenwerfer, H. (1985). *Monotonie in unserer Arbeitswelt – muß das sein* [Is it necessary to have monotony in our industry]*?* Muenchen: Bayrisches Staatsministerium fuer Arbeit und Sozialordnung.

25. Baschera, P. & Greandjean, E. (1979). Effects of repetitive tasks with different degrees of difficulty on critical fusion frequency and subjective state. *Ergonomics*, 22, 377-385.

26. Battiste, V. & Bortolussi, M. (1988). *Transport pilot workload: A comparison of two subjective techniques.* In Proceedings of the Human Factors Society Thirty-Second Annual Meeting (pp. 150-154). Santa Monica, CA: Human Factors Society.

27. Bauer, H., Guttmann, G., Leodolter, M., & Leodolter, U. (2003) *Test zur Zeit- und Bewegungsantizipation* (ZBA) [Time-Movement-Anticipation]. Moedling: Schuhfried.

28. Bauer, P. (1991). Multiple testing in clinical trials. *Statistics in Medicine*, 10, 871-890.

29. Bauer, P., Roehmel, J., Maurer, W., & Hothorn, L.A .(1998). Testing Strategies in Multiple-Dose Experiments Including Active Control. *Statistics in Medicine*, 17, 2133-2146.

30. Baumgartner, M. (2004). Critical incident stress management in air traffic control. EEC Note No. 15/04. EUROCONTROL: Bretigny sur Orge.

31. Bedny, G. Z. & Karwowski, W. (2004). Activity theory as a basis for the study of work. *Ergonomics*, 47(2), 134-154.

32. Belyavin, A. & Wright, N.A. (1987) Changes in electrical activity of the brain with vigilance. *Electroencephalo-graphy and Clinical Neurophysiology*, 66, 137-144.

33. Berlyne, D. (1960). *Conflict, Arousal and couriosity*. New Yor: McGraw-Hill.

34. Berman, A. (1939a). The effect of benzendrine sulfate (amphetamine sulfate) on satiation. *American Journal of Psychology*, 52, 297-299.

35. Berman, A. (1939b). The relation of time estimation to satiation. *Journal of Experimental Psychology*, 25, 281-293.

36. Berman, A. (1939c). Satiation or boredom? *American Journal of Psychology*, 52, 471-473.

37. Bernstein, H.E. (1975). Boredom and the ready-made life. *Social Research*, 42, 512-537.

38. Berntson, G. G., Bigger, J. T., Jr., Eckberg, D. L., Grossman, P., Kaufmann, P. G., Malik, M., et al. (1997). Heart rate variability: origins, methods, and interpretive caveats. *Psychophysiology*, 34(6), 623-648.

39. Berntson, G. G., Cacioppo, J. T., & Quigley, K. S. (1993). Respiratory sinus arrhythmia: autonomic origins, physiological mechanisms, and psychophysiological implications. *Psychophysiology*, 30(2), 183-196.

40. Bevan, W., Avant, L., & Lankford, H. G. (1967). Influence of interpolated periods of activity and inactivity upon the vigilance decrement. *Journal of Applied Psychology*, 51(4), 352-356.

41. Bingham, C. R., Stemmler, M., Petersen, A. C., & Graber, J. A. (1998). Imputing missing data values in repeated measurement within-subjects designs. *Methods of Psychological Research*, 3(2), 131-155.

42. Boehm-Davis, D. A., Gray, W. D. & Schoelles, M. J. (2000). The eye blink as a physiological indicator of cognitive workload. *Proceedings of the IEA 2000/HFES 2000 Congress Volume 6* (pp. 116-119). Santa Monica, CA: Human Factors and Ergonomics Society.

43. Boquet, A., Cruz, C. E., Nesthus, T. E., Detwiler, C. A., Knecht, W. R., & Holcomb, K. A. (2002). *A Laboratory Comparison of Clockwise and Counter-Clockwise Rapidly Rotating Shift Schedules*, Part III: Effects on Core Body Temperature and Neuroendocrine Measures (DOT/FAA/AM-02/20). Washington, DC: Department of Transportation/Federal Aviation Administration, Office of Aviation Medicine.

44. Boring, E.G., (1957). *A history of experimental psychology.* (2nd ed.). New York: Appleton-Century-Crofts.

45. Bortz, J. (1999) Statistik fuer Sozialwissenschaftler [Statistics for Social Scientists]. (5th ed.). Berlin: Springer.

46. Bossong, B. (1999) *Stress und Handlungskontrolle. Die volitionale Kompetenz bei der Bewertung und Bewaeltigung aversiver Erfahrungen* [Stress and action control]. Band 17 der Reihe Motivationsforschung. Goettingen: Hogrefe.

47. Boucsein, W. (1988). *Elektrodermale Aktivitaet [Electrodermal Activity]*. Berlin: Springer.

48. Boucsein, W. (1991). Arbeitspsychologische Beanspruchungsforschung heute – eine Herausforderung an die Psychophysiologie [Research on Mental workload – a challenge for psychophysiology]. *Psychologische Rundschau*, 42, 129-144.

49. Boucsein, W. (2000). The Use of Psychophysiology for Evaluating Stress-Strain Processes in Human-Computer Interaction. In R. W. Backs & W. Boucsein (Eds.). *Engineering psychophysiology. Issues & Applications* (pp. 289-309). London: Lawrence Erlbaum Associates.

50. Boucsein, W. & Backs, R. W. (2000). Engineering psychophysiology as a Discipline: Historical and Theoretical Aspects. In R. W. Backs & W. Boucsein (Eds.). *Engineering psychophysiology. Issues & Applications* (pp. 3-30). London: Lawrence Erlbaum Associates.

51. Boucsein, W. & Thum, M. (1997). Design of work/rest schedules for computer work based on psychophysiological recovery measures. *International Journal of Industrial Ergonomics*, 20, 51-57.

52. Braby, C. D., Harris, D., & Mur, H. C. (1993). A psychophysiological approach to the assessment of work underload. *Ergonomics*, 36 (9), 1035-1042.

53. Branton, P. (1970). A field study of repetitive manual work in relation to accidents at the work place. *International Journal of Production Research*, 8, 93-107.

54. Briner, R. B. (1996). The neglect and importance of emotions at work. *European Journal of Organizational Psychology*, 8(3), 323-346.

55. Brocke, B., Tasche, K. G., & Beauducel, A. (1997). Biopsychological foundations of extraversion: Differential effort reactivity and state control. *Personality and Individual Differences*, *22*, 447-458.

56. Brookhuis, K. A. & De Waard, D. (2002). On the assessment of (mental) workload and other subjective qualifications. *Ergonomics*, 45(14), 1026-1030.

57. Brookings, J. B., Wilson, G. F., & Swain, C. R. (1996). Psychophysiological responses to changes in workload during simulated air traffic control. *Biological Psychology*, 42, 361-377.

58. Buunk, B. P., De Jonge, J., Ybema, J. F., & De Wolff, C. J. (1998). In P.J.D. Drenth, H. Thierry & C.J. De Wolff (Eds.) *Work psychology. Handbook of Work and Organizational Psychology* (2nd ed.) (pp. 145-182). Hove: Psychology Press.

59. Byers, J. C., Bittner, A. C., & Hill, S. G. (1989). Traditional and raw Task Load Index correlations: Are paired comparisons necessary? In A. Mital (Ed.) *Advances in industrial ergonomics and safety* (pp. 481-485). London: Taylor & Francis.

60. Byrne, E. A. & Parasuraman, R. (1996). Psychophysiology and adaptive automation. *Biological Psychology*, 42, 249-268.

61. Cacioppo J., Petty R., & Kao, C. (1984). The efficient assessment of need for cognition. *Journal of Personality Assessment*, 48, 306-307.

62. Cacioppo, J. T. & Tassinary, L. G. (1990). Inferring Psychological Significance From Physiological Signals. *American Psychologist*, 45(1), 16-28.

63. Cacioppo, J. T., Tassinary L. G, & Berntson, G. G. (2000). Psychophysiological Science. In J. T. Cacioppo, L. G. Tassinary and G. G. Berntson (Eds.) *Handbook of Psychophysiology* (2nd ed.) (pp. 3-23). Cambridge: University Press.

64. Caldwell, J. A. & Ramspott, S. (1998). Effects of task duration on sensitivity to sleep deprivation using the multi-attribute task battery. *Behavior Research Methods, Instruments and Computers*. 30 (4), 651-660 [Abstract].

65. Cattell, R. B. (1966). *Handbook of multivariate experimental psychology.* Chicago: Rand McNally.

66. Cohen, J. (1992). Statistical Power Analysis for the Behavioural Sciences (2nd ed.). Hillsdale, NJ: Erlbaum.

67. Collet, C., Averty, P., Delhomme, G., Dittmar, A., & Vernet-Maury, E. (2003). Subjective aspects of mental workload in air traffic control. In G. R. J. Hockey, A. W. K. Gaillard, & O. Burov (Eds.). *Operator Functional State* (pp. 291-302). Amsterdam: IOS Press.

68. Costa, G. (1993). Evaluation of workload in air traffic controllers. *Ergonomics*, 36 (9), 1111-1120.

69. Costa, G. (1999). Fatigue and Biological Rhythms. In D. J. Garland (Ed.). *Handbook of aviation human factors. (*pp. 235-254). Mahwah, NJ: Erlbaum.

70. Costa, P. T., Jr., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI): Professional Manual.* Odessa, FL: Psychological Assessment Resources.

71. Cox, T. (1985). Repetitive work: Occupational stress and health. In C. L. Cooper & M. J. Smith (Eds.). *Job stress and blue collar work* (pp. 85-112). Wiley: Chichester.

72. Crabbe, J. B. & Dishman, R. K. (2004). Brain electrocortical activity during and after exercise: A quantitative synthesis. *Psychophysiology*, 41, 563-574.

73. Csikszentmihalyi, M. (1975). *Beyond Boredom and Anxiety.* San Francisco: Jossey-Bass.

74. Csikszentmihalyi, M. & LeFevre, J. (1989). Optimal experience in work and leisure. *Journal of Personality and Social Psychology*, 56(5), 815-822.

75. David, H., Mollard, R., Cabon, P. & Farbos, B. (2000). Point-of-Gaze Evaluation in Simulated ATC. *Proceedings of the IEA 2000/HFES 2000 Congress Volume 6*, Santa Monica, CA: Human Factors and Ergonomics Society, 666-669.

76. Davies, D. R., Shackleton, V.J., & Parasuraman, R. (1983). Monotony and Boredom. In G.R.J. Hockey (Ed.) *Stress and Fatigue in Human Performance*. New York: Wiley.

77. Deaton, J.E. & Parasuraman, R. (1993). Sensory and cognitive vigilance: Effects of age on performance and mental workload. *Human Performance*, 6, 71-97.

78. Della Rocco, P. S. & Cruz, C. E. (1999). Operational Errors/Deviations and Shiftwork in Air Traffic Control. In P. S. Della Rocco, (Ed). *The Role of Shift Work and Fatigue in Air Traffic Control Operational Errors and Incidents* (pp.11-21). (DOT/FAA/AM-99/02). Washington, DC: Department of Transportation/Federal Aviation Administration, Office of Aviation Medicine.

79. Demerouti, E., Bakker, A. B., Nachreiner, F. & Ebbinghaus, M. (2002). From mental strain to burnout. *Journal of work and organizational psychology*, 11(4), 423-441.

80. Desmond, P. & Hancock, P. (2001). Active and passive fatigue states. In P. Hancock (Ed.). *Stress, workload and fatigue* (pp. 455-465). Mahwah: Erlbaum.

81. Desmond, P. A. & Hoyes, T. W. (1996). Workload variation, intrinsic risk and utility in a simulated air traffic control task: evidence for compensatory effects. *Safety Science*, 22(1-3), 87-101.

82. Desmond, P. A., Matthews, G., & Bush, J. (2001). Sustained visual attention during simultaneous and successive vigilance tasks. *Proceedings of the Human Factors and Ergonomics Society 45th Annual Meeting (pp.* 1386-1389). Santa Monica, CA: Human Factors and Ergonomics Society.

83. Dey, A. K., Abowd, G. D., & Salber, D. A. (2001). A conceptual framework and a toolkit for supporting the rapid prototyping of context-aware applications. *Human-Computer Interaction*, 16, 97-166.

84. Dickman, S.J. (2002). Dimensions of Arousal: Wakefulness and Vigor. *Human Factors*, 44(3), 429-442.

85. Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology, 41*, 417-440.

86. Dinges, D. F., Pack, F., Williams, K., Gillen, K. A., Powell, J. W., Ott, G. E., Aptowicz, C., & Pack, A. (1997). Cumulative sleepiness, mood disturbance, and psychomotor vigilance performance decrements during a week of sleep restricted to 4-5 hours per night. *Sleep*, 20(4), 267-277.

87. Dougherty, M., Gronlund, S., & Durso, F. (1999). Plan generation in air traffic control. In R. Jensen et al. (Eds.). *10th International Symposium on Aviation Psychology* (pp. 541-547). The Ohio State University.

88. Eberspaecher, H., Hermann, H., & Kallus, K. W. (1993). Psychische Erholung und Regeneration zwischen Beanspruchungen [Recovery and Regeneration between strain]. In J. R. Nitsch & R. Seiler (Eds.), *Bewegung und Sport – Psychologische Grundlagen und Wirkungen, Bd. 1 (pp. 237-241)*. Sankt Augustin: Academia Verlag.

89. Edkins, G. D. & Pollock, C. M. (1997). The influence of sustained attention on Railway accidents. *Accident Analysis & Prevention*, 29, 4, 533-539.

90. Eilers, K., Nachreiner, F., & Boening E. (1989). Zur subjektiven Skalierung psychischer Beanspruchung [Subjective ratings of strain]. *Zeitschrift fuer Arbeitswissenschaft*, 44(1), 217-223.

91. Erdfelder, E. (1994). Erzeugung und Verwendung empirischer Daten [The generation and use of empirical data]. In T. Herrmann & W. H. Tack (Eds.). *Methodologische Grundlagen der Psychologie*, Enzyklopaedie der Psychologie, Themenbereich B, Serie I, (Vol 1, pp. 47-97). Goettingen: Hogrefe.

92. Erdfelder, E., Faul, F., & Buchner, A. (1996). GPOWER: A general power analysis program. *Behavior Research Methods, Instruments, & Computers, 28*, 1-11.

93. EUROCONTROL (2005). *COSAAC User Manual.* EUROCONTROL Experimental Centre: Bretigny sur Orge.

94. Fahrenberg, J., Foerster, F., Schneider, J. J., Mueller, W., & Myrtek, M. (1984). *Aktivierungsforschung im Labor-Feld-Vergleich* [Comparing activation research in the lab and field]. Muenchen: Minerva.

95. Fahrenberg, J., Walschburger, P., Foerster, F., Myrtek, M., & W. Mueller (1979). *Psychophysiologische Aktivierungsforschung.* Ein Beitrag zu den Grundlagen der multivariaten Emotions- und Stress-Theorie [Psychophysiological activation research]. Muenchen: Minerva.

96. Fahrenberg, J. & Wientjes, C. J. (2000). Recording Methods in Applied Environments. In R. W. Backs & W. Boucsein (Eds.), *Engineering psychophysiology. Issues & Applications* (pp. 111 - 138). London: Lawrence Erlbaum Associates.

97. Farmer, R. & Sundberg, N.D. (1986). Boredom proneness: the development and correlates of a new scale. *Journal of personality assessment*, 50(1), 4-17.

98. Fenichel, O. (1951). On the psychology of boredom. In D. Rapaport (Ed.) *Organization and pathology of thought (pp.* 349 – 361). New York: Columbia University Press.

99. Finkelman, J. M. (1994). A large database study of the factors associated with work-induced fatigue. *Human factors*, 36 (2), 232-243.

100. Fisher, B. C. (1998). *Attention Deficit Disorder Misdiagnosis.* Washington D.C.: CRC Press.

101. Fisher, C. (1998). Effects of external and internal interruptions on boredom at work: two studies. *Journal of Organizational Behavior*, 19, 503-522.

102. Foerster, F., Myrtek, M., & Stemmler, G. (1993). Reactivity to multiple stressors: A course in synergism. *Journal of Psychophysiology*, 7, 115-124.

103. Fowles, D. C. (1980). The three arousal model: Implications of Gray's two factor learning theory for heart rate, electrodermal activity and psychopathy. *Psychophysiology*, *17*, 87-104.

104. Frankenhauser , M. A. (1986). Psychobiological Framework for Research on Human Stress and Coping. In M. H Appley & R. Trumbull (Eds.). *Dynamics of Stress: Physiological, Psychological and Social Perspectives* (pp. 101-116). New York: Plenum Press.

105. Frankenhaeuser, M. (1971a). Behavior and circulating catecholamines. *Brain Research*, 31(2), 241-262.

106. Frankenhaeuser, M. (1971b). Physiological Reactions to understimulation and overstimulation. *Acta Psychologica*, 35, 298-308.

107. Friedland, D. S. & Price, R. H. (2003). Underemployment: Consequences for the Health and Well-Being of Workers. *American Journal for Community Psychology*, 32(1-2), 33-45.

108. Frishman, E. Y. (1990). The dynamics of human sensory-perceptive activity in conditions of monotony and fatigue: The psychophysical approach. *Soviet Journal of Psychology*, 11(5), 21-32.

109. Gaillard, A. W. K. (2003). Fatigue Assessment and Performance Protection. In G. R. J. Hockey, A. W. K. Gaillard, & O. Burov (Eds.). *Operator Functional State* (pp. 24-35). Amsterdam: IOS Press.

110. Gaillard, A. W. K. (2005). Concentration – an instrument to augment cognition. *Proceedings of the 1st International Conference on Augmented Cognition*, Las Vegas, NV, 22-27 July 2005.

111. Gaillard, A. W. K., & Wientjes, C. J. E. (1994). Mental load and work stress as two types of energy mobilization. *Work Stress*, 8, 141-152.

112. Galley, N. (1993). The evaluation of the electrooculogram as a psychophysiological measuring instrument in the driver study of driver behaviour. *Ergonomics*, 36(9), 1063-1070.

113. Galley, N. (1998). An enquiry into the relationship between activation and performance using saccadic eye movement parameters. *Ergonomics*, 41(5), 698-720.

114. Gander, P. (2001). Fatigue management in Air Traffic Control: The New Zealand Approach. *Transportation Research Part F4*, 49-62.

115. Gawron, V. J. (2000). *Human Performance Measures Handbook.* Mahwah: Lawrence Erlbaum Associates.

116. Gawron, V. J., French, J., & Funke, D. (2001). An overview of fatigue. In P. A. Hancock & P. Desmond (Eds.). *Stress, workload and fatigue* (pp.581-594). Mahwah: Erlbaum.

117. Geiwitz, P. J. (1966). Structure of Boredom. *Journal of Personality and Social Psychology*, 3(5), 592-600.

118. Gendolla, G. H. E. & Krueksen, J. (2001). The joint impact of mood state and task difficulty on cardiovascular and electrodermal reactivity in active coping. *Psychophysiology*, 38, 548-556.

119. Gereb, G. (1968). Some psychological problems of monotony and stress. *Studia Psychologica*, 10(4), 313-324.

120. Gereb, G. (1978). Experiments from the domain of monotony and vigilance. *Studia Psychologica,* 20(1), 35-45.

121. Gillberg, M.& Åkerstedt, T. (1998). Sleep loss and performance: no "safe" duration of a monotonous task, *Physiology & Behavior*, 64(5), 599-604.

122. Gillberg, M., Kecklund, G., Goeransson, B., & Akerstedt, T. (2001). Operator Performance and signs of sleepiness during day and night work in simulated thermal power plant. *International Journal of Fatigue*, 23(9), 751-766.

123. Goldberg, L. R. (1990). An alternative "Description of Personality": The Big-Five Factor Structure. *Journal of Personality and Social Psychology*, 59, 1216-1229.

124. Goldberg, L. R. (1999). The Development of Five-Factor Domain Scales from the IPIP Item Pool. [On-line]. Available URL: http://ipip.ori.org/ipip/memo.htm.

125. Gopher, D., & Donchin, E. (1986). Workload - An Examination of the Concept. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of Perception and Human Performance* (pp. 41-1 - 41-49). New York, NY: John Wiley and Sons.

126. Graf, O. (1961). Arbeitszeit und Arbeitspausen [Work Schedules and Rest Breaks]. In A. Mayer & B. Herwig (Eds.). *Handbuch der Psychologie,* Betriebspsychologie (Vol 9, pp. 95-117). Goettingen: Hogrefe.

127. Grandjean, E. (1991). Physiologische Arbeitsgestaltung. Leitfaden der Ergonomie [Physiological work design - Guidelines for Ergonomy] (4th revised ed.). Thun: Ott.

128. Gratton, G. (2000). Biosignal Processing. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of Psychophysiology* (2nd ed.) (pp. 900-923). UK: Cambridge.

129. Greenfield, N. S. & Sternbach, R. A. (1972). *Handbook of psychophysiology*. New York: Holt, Rinehart and Winston.

130. Gregory, K. B., Oyung, P. J. & Rosekind, M. R. (1999). Managing alertness and performance in Air Traffic Control Operations. In P. S. Della Rocco (Ed). The Role of Shift Work and Fatigue in Air Traffic Control Operational Errors and Incidents (pp. 1-10). (DOT/FAA/AM-99/02). Washington, DC: Department of Transportation/ Federal Aviation Administration, Office of Aviation Medicine.

131. Greif, S. (1991). Stress in der Arbeit – Einfuehrung und Grundbegriffe. In S. Greif, E. Bamberg & N. Semmer (Eds.), *Psychischer Stress am Arbeitsplatz* (pp. 1-28). Goettingen: Hogrefe.

132. Groß, B. & Metz, A. M. (1998). Evozierte Hirnpotentiale als Indikatoren fuer Arbeitsbeanspruchungen [Event-related potentials as indicators for mental workload]. *Zeitschrift fuer Arbeits- und Organisationspsychologie*, 42(4), 213-218.

133. Gubser, A. (1968). *Monotonie im Industriebetrieb [Monotony in Industry]*. Bern: Hans Huber.

134. Guttman, J. A. & Stein, E. S. (1997). *En route generic airspace evaluation* (DOT/FAA/CT-TN97/7). Atlantic City International Airport: Federal Aviation Administration, William J. Hughes Technical Center.

135. Hacker, W. (1982). Beanspruchungskomponenten von geistigen Routinetaetigkeiten [Strain components of mental routine activities]. *Zeitschrift fuer Psychologie*, 190, 233-258.

136. Hacker, W. (2003). Action Regulation Theory: A practical tool for the design of modern work processes? *European Journal of work and organizational psychology*, 12(2), 105-130.

137. Hacker, W. & Richter, P. (1984). *Psychische Fehlbeanspruchung. Psychische Ermuedung, Monotonie, Saettigung und Streß [Psychic strain: fatigue, monotony, satiation and stress]* (2nd revised ed.). Berlin: Springer

138. Haga, S., Shinoda, H. & Kokubun, M. (2002). Effects of task difficulty and time-on-task on mental workload. *Japanese Psychological Research*, 44(3), 134-143.

139. Hagemann, T. (2000). *Belastung, Beanspruchung Und Vigilanz in Der Flugsicherung: Unter Besonderer Berucksichtigung Der Towerlotsentatigkeit [Load, strain and vigilance in ATC].* Europaeische Hochschulschriften, Reihe 6, Psychologie, 664, Frankfurt am Main: Lang.

140. Hagenmeyer, L. (2005). *Preliminary guidelines for the personalization of the HMI of hypovigilance warning systems.* Paper presented at HCI 2005, Las Vegas, Nevada, USA, 22-27 July 2005.

141. Haider, M. (1956). Experimentelle Untersuchungen ueber Daueraufmerksamkeit und cerebrale Vigilanz bei einfoermigen Taetigkeiten [Experiments on sustained attention and vigilance in uniform tasks]. *Zeitschrift für experimentelle und angewandte Psychologie*, 10(1), 1-18.

142. Hancock, P. A. & Warm, J. S. (1989). A dynamic model of stress and sustained attention. *Human Factors, 31,* 519-537.

143. Hancock, P. A., Williams, G., Miyake, S., Manning, C. M. (1995). Influence of Task Demand Characteristics on Workload and Performance. *International Journal of Aviation Psychology*, 5, 63-86.

144. Harris, M. B. (2000). Correlates and characteristics of boredom proneness and boredom. *Journal of Applied Social Psychology*, 30, 576-598.

145. Hart, S. G. & Staveland, L. E. (1988, reprint 2004). The development of the NASA-TLX. In N. Moray, (Ed.). *Ergonomics: Major Writings.* Volume II. Skill, Displays, Controls and Mental Workload (pp. 408-463). London: Taylor & Francis.

146. Hart, S. G. & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index). Results of empirical and theoretical research. In P.A. Hancock & N. Meshkati (Eds.). *Human Mental Worklad* (pp. 139-183). Amsterdam: North-Holland.

147. Head, H. (1923). The conception of nervous and mental energy. *The British Journal of Psychology. General Section*, 14, 126-147.

148. Hennig, J., Kieferdorf, P., Moritz, C., Huwe, S. & Netter, P. (1998) Changes in cortisol secretion during shiftwork: implications for tolerance to shiftwork? *Ergonomics*, 41(5), 610-621).

149. Heslegrave, R. J., Ogilvie, J. C. & Furedy, J. J. (1979). Meas,uring Baseline-Treatment Differences Mean Square and Beats per Minute versus Interbeat Intervals. *Psychophysiology*, 16(2), 151-157.

150. Higuchi, S., Liu, Y., Yuasa, T., Maeda, A., & Motohashi, Y. (2001). Diurnal variations in alpha power density and subjective sleepiness while performing repeated vigilance tasks. *Clinical Neurophysiology*, 112(6), 997-1000.

151. Hill, A. B. (1975). Work variety and Individual Differences in Occupational Boredom. Journal *of Applied Psychology*, 60(1), 128-131.

152. Hill, A. B. & Perkins, R. E. (1985). Towards a model of boredom. *British Journal of Psychology*, 76, 235-240.

153. Hilburn, B. & Jorna, P. (2001). Workload and air traffic control. In P. A. Hancock & P. Desmond (Eds.). *Stress, workload and fatigue* (pp. 384-394). Mahwah: Erlbaum.

154. Hilburn, B. G., Bakker, M. W. P., Pekela, W. D., & Parasuraman R. (1997). *The Effect of Free Flight on Air Traffic Controller Mental Workload, Monitoring and System Performance.* Paper presented at the CEAS Symposium on 'Free Flight', Amsterdam.

155. Hinz, A., Seibt, R., Hueber, B., & Schreinicke, G. (2000). Response Specificity in Psychophysiology: A Comparison of Different Approaches. *Journal of Psychophysiology*, 14(2), 115-122.

156. Hitchcock, E. M., Dember, W. N., Warm, J. S., Moroney, B. W., & See, J. E. (1999). Effects of Cueing and Knowledge of results on Workload and Boredom in Sustained Attention. *Human Factors*, 41(3), 365-372.

157. Hobfoll, S. E. (1998). *Stress, Culture, and Community: The Psychology and Philosophy of Stress.* New York: Springer.

158. Hockey, G. R. J. (1986). A state control theory of adaptation and individual differences in stress management. In G. R. J. Hockey, A. W. K. Gaillard, & M. G. H. Coles (Eds.), Energetics and Human Information Processing (pp. 285-298). Dordrecht: Kluwer Academic.

159. Hockey, G. R. J. (1997) Compensatory control in the regulation of human performance under stress and high workload: A cognitive-energetical framework. *Biological Psychology*, 45, 73-93.

160. Hockey, G. R. J. (2003). Operator Functional State as a Framework for the assessment of Performance Degradation. In G. R. J. Hockey, A. W. K. Gaillard, & O. Burov (Eds.). *Operator Functional State* (pp.3-7). Amsterdam: IOS Press.

161. Hockey, G. R. J., Gaillard, A. W. K., & Coles, M. G. H. (1986). *Energetics and human information processing*. Dordrecht: Matinus Nijhoff.

162. Hockey, G. R. J. & Tattersall, A. J. (1989). The maintenance of vigilance during automated monitoring. In A. Coblentz (Ed.) *Vigilance and Performance in Automatized Systems/Vigilance et Performance de L'Homme dans les Systemes Automatises* (pp. 13-22). The Netherlands: Kuwer Academic.

163. Hockey, G. R. J., Wastell, G., & Sauer, J. (1998). Effects of sleep deprivation and user interface on complex performance: A multilevel analysis of compensatory control. *Human Factors*, 40(2), 233-253.

164. Hohnsbein, J., Falkenstein, M., & Hoormann J. (1998). Performance differences in reaction tasks are reflected in event-related brain potentials (ERPs). *Ergonomics*, 41(5), 622-633.

165. Hohnsbein, J., Falkenstein, M., & Hoormann, J. (1995). Effects of attention and time-pressure subcomponents and implications for workload research on P300. *Biological Psychology*, 40, 73-81.

166. Horne, J. A. & Ostberg, O. (1976). A Self Assessment Questionnaire to Determine Morningness-Eveningness in Human Circadian Rhythms. *International Journal of Chronobiology*, 4, 97-110.

167. Hot, P., Naveteur, J., Leconte, P., & Sequeira H. (1999). Diurnal variations of tonic electrodermal activity. *International Journal of Psychophysiology*, 33(3), 223-230.

168. Hukki, K. & Norros, L. (1990) Subject-centered and systematic conceptualization as a tool of simulator training. *Le travail human*, 61(4),313-331.

169. Hulin, C. L. & Blood, M. R. (1968). Job enlargement, individual differences and worker responses. *Psychological Bulleting*, 69, 41-55.

170. Idogawa, K. (1991). On the brain waves activity of professional drivers during monotonous work. *Behaviormetrika*, 30, 23-34.

171. ISO 10075 Ergonomic principles related to mental workload, Part 1: ISO 10075-1:1991 General terms and definitions, Part 2: ISO 10075-2:1996 Design principles, Part 3: ISO 10075-3: 2004 Principles and requirements concerning methods for measuring and assessing mental workload. International Organization for Standardization, Genf.

172. Jain, A., Martens, W., Mutz, G., Weiss, R. W. & Stephan, E. (1996). Towards a comprehensive technology for recording and analysis of multiple physiological parameters within their behavioral and environmental context. In J. Fahrenberg & M. Myrtek (Eds.). *Ambulatory Assessment* (pp. 215-235). Seattle: Hogrefe & Huber Publishers.

173. Jain, A., Schmidt T. F. H., Johnston, D. W., Brabant G., & Von zur Muehlen, A. (1998). The relationship between heart rate and blood pressure reactivity in the laboratory and in the field: Evidence using continuous measures of blood pressure, heart rate and physical activity *Journal of Psychophysiology*, 12(4).

174. Janke, W. (1976). Fragebogen zur Erfassung der Ausgangslage (FAL) [The initial State questionnaire]. Wuerzburg: University, Institute for Psychology.

175. Jansen N. W., Van Amelsvoort L. G., Kristensen T. S., Van den Brandt, P. A., & Kant I. J. (2003). Work schedules and fatigue: a prospective cohort study. *Occupational and Environmental Medicine*, 60(1), 47-53.

176. Janssen N., Kant I. J., Swaen G. M., Janssen P. P., & Schroer, C. A. (2003). Fatigue as a predictor of sickness absence: results from the Maastricht cohort study on fatigue at work. *Occupational and Environmental Medicine*, 60(1), 71-76.

177. Jeannot, E., Kelly, C. & Thompson, D. (2003). *The Development of Situation Awareness Measures in ATM Systems* (HRS/HSP-005-REP-01). Brussels: EUROCONTROL.

178. Jennings, J. R., Kamarck, T., Stewart, C., Eddy, M., & Johnson, P. (1992). Alternate cardiovascular baseline assessment techniques: Vanilla or resting baseline. *Psychophysiology*, 29, (6), 742-750.

179. Jimenez (2000). Profilanalyse der Arbeitszufriedenheit (PAZ) – Test Manual [Work Satisfaction Profile]. Eggers-dorf: Research Team.

180. Johansson, G. (1989). Job demands and stress reactions in repetitive and uneventful monotony at work. *International Journal of Health Services,* 19(2), 365-377.

181. Johansson, G., Cavalini, P. M., & Petterson, P. (1996). Psychobiological reactions to unpredictable performance stress in a monotonous situation. *Human Performance*, 9(4), 363-384.

182. Johansson, G. & Sanden, P. O. (1989). *Mental Load and job satisfaction of control room operators*. Reports from the Department of Psychology, University of Stockholm, No. 698.

183. Jorna, P. G. (1993). Heart rate and workload variations in actual and simulated flight. *Ergonomics*, 36 (9), 1043-1054.

184. Kahneman, D. (1973). Attention and effort. Englewood Cliffs, NJ: Prentice Hall.

185. Kallus, K. W. (1991). Initial State and the Law of Initial Value. *Journal of Psychophysiology*, 5, 109-110.

186. Kallus, K. W. (1992). *Beanspruchung und Ausgangszustand* [Strain and Initial State]. Weinheim: Psychologie Verlags Union.

187. Kallus, K. W. (1995). Der Erholungs-Belastungs-Fragebogen (EBF) [The Recovery-Stress-Questionnaire]. Frankfurt a.M.: Swets Test Servises.

188. Kallus, K. W., Barbarino, M., & Van Damme, D. (1997). *Model of the Cognitive Aspects of Air Traffic Control.* EUROCONTROL (HUM.ET1.ST01.1000.REP-02). Brussels: EUROCONTROL.

189. Kallus, K. W., Barbarino, M., Van Damme, D., & Dittman, A. (1999). Integrated Task Analysis of En-Route Control: A process oriented approach. In R. S. Jensen, B. Cox, D. Callister & R. Lavis (Eds.). *Proceedings of the 10th International Symposium on Aviation Psychology* (pp.517-521). Columbus: The Ohio State University.

190. Kallus, K. W., Hoffmann, P., Ehgartner, B., Kuhn, C., Pichler, A., & Schuen-Medwed, R. (2003). Safety-Efficiency-Workload Balance in ATC: A tool to assess sector capacity from a human factors perspective. *Proceedings of the 12th International Symposium on Aviation Psychology* (pp. 639-643). Wright State University Dayton, Ohio.

191. Karasek, R. & Theorell, T. (1990). Healthy work: stress, productivity, and the reconstruction of working life. New York: Basic Books.Abt, K. (1983). Significance testing of many variables. *Neuropsychobiology*, 9, 47-51.

192. Karsten, A. (1928). Psychische Saettigung [Psychic Satiation]. In K. Lewin (Ed.), *Psychologische Forschung*, 10, 142-254.

193. Kass, S. J., Vodanovich, S. J., & Callender, A. (2001). State-trait boredom: relationship to absenteeism, tenure, and job satisfaction. *Journal of Business and Psychology*, 16, 317-327.

194. Kass, S. J., Vodanovich, S. J., Stanny, C., & Taylor, T. (2001). Watching the clock: Boredom and vigilance performance. *Perceptual and Motor Skills*, 92, 969-976.

195. Koelega, H.S. (1992). Extraversion and Vigilance Performance: 30 Years of Inconsistencies. *Psychological Bulletin*, 112, 239-258.

196. Koella, W. P. (1982). A modern neurobiological concept of vigilance. *Experientia*, 38, 1426-1437.

197. Kogi, K. (1972). Repeated short indifference period in industrial vigilance. *Journal of Human Ergology*, 1, 111-121.

198. Kohlisch, O., Kuhmann, W., & Boucsein, W. (1991). Auswirkungen systembedingter Arbeitsunterbrechungen bei computergestützter Textverarbeitung – Eine Feldstudie. *Zeitschrift für Experimentelle und Angewandte Psychologie, 38, 585-604*.

199. Kok, A. (1997). Event-related-potential (ERP) reflections of mental resources: a review and synthesis. *Biological Psychology*, 45, 19-56.

200. Kopardekar, P. & Magyarits, S. (2003). *Measurement and Prediction of Dynamic Density*. Paper presented at ATM 2003, FAA EUROCONTROL Meeting.

201. Kozena, L., Frantik, E., & Dvorak, J. (1996) Vigilance enhanced by psychophysiological means: Comparison of several methods. *Homeostasis in Health and Disease*, 37(6), 256-260.

202. Kraeplin, E. (1903). Ue*ber Ermuedungsmessungen.* Archiv ges. Psychologie, 1.

203. Krulewitz, J. E., Warm, J. S. & Wohl, T. H. (1975). Effects of shifts in the rate of repetitive stimulation on sustained attention. *Perception and Psychophysics*, 18, 245-249.

204. Kuhl, J. (1994a). A theory of action and state orientations. In J. Kuhl & J. Beckmann (Eds.), *Volition and personality: Action versus state orientation* (pp. 9-46). Seattle, WA: Hogrefe & Huber.

205. Kuhl, J. (1994b). Action versus state orientation: Psychometric properties of the Action Control Scale (ACS-90). In J. Kuhl & J. Beckmann (Eds.), *Volition and personality: Action versus state orientation* (pp.47-60). Seattle, WA: Hogrefe & Huber.

206. Kuiper, J. I., Van der Beek, A. J., & Meijman, T. F. (1998). Psychosomatic complaints and unwinding of sympathoadrenal activation after work. *Stress and Health*, 14(1), 7-12.

207. Lacey, J. L. (1967). Somatic response patterning and stress: some revisions of activation theory. In M. H. Appley & R. Trumbell (Eds.), *Psychological Stress: Issues in research* (pp.14-42). New York: Appleton-Century-Croft.

208. Lane, A. M., & Lovejoy, D. J., (2001). The effects of exercise on mood changes: The moderating effect of depressed mood. *Journal of sports medicine and physical fitness*, 41 (4), 539-548.

209. Larsen, R. J., & Diener, E. (1992). Problems and promises with the circumplex model of emotion. *Review of Personality and Social Psychology*, 13, 25-59.

210. Laudeman, I. V., Shelden, S. G., Branstrom, R. & Brasil, C.L. (1998*). Dynamic Density. An Air Traffic Management Metric.* California: National Aeronautics and Space Administration, Ames Research Center, NASA/TM-1998-112226.

211. Lazarus, R.S. (1993). From psychological stress to the emotions: A history of changing outlooks. *Annual Review of Psychology*, 44, 1-21.

212. Lazarus, R. S. & Folkman, S. (1984). *Stress, appraisal, and coping.* New York: Springer.

213. Leclerq, M. (2002). Theoretical aspects of the main components and functions of attention. In M. Leclercq & P. Zimmermann (Eds.), *Applied Neuropsychology of Attention* (pp. 3-55). New York: Psychology Press.

214. Leonova, A. (2003). Functional Status and Regulatory Processes in Stress Management. In G. R. J. Hockey, A. W. K. Gaillard, & O. Burov (Eds.). *Operator Functional State* (pp.36-52). Amsterdam: IOS Press.

215. Leontjew, A. N. (1982*). Taetigkeit , Bewusstsein, Persoenlichkeit* . Koeln: Pahl-Rugenstein.

216. Lille F. & Cheliout, F. (1982). Variations in diurnal and nocturnal waking state in air traffic controllers. *European Journal of Applied Physiology*, 49, 319-328.

217. Lindsay, M. & Ehrenberg, A. S. C. (1993). The design of replicated studies. *The American statistician*, 47(3), 217-228.

218. Linnenbrink, E. A. & Pintrich, P. R. (2004). Role of affect in cognitive pro-cessing in academic contexts. In D.Y. Dai & R. Sternberg. *Motivation, emotion, and cognition: integrative perspectives on intellectual development and functioning* (pp. 57-88). Mahwah, NJ: Lawrence Erlbaum Associates.

219. Loehr, G. & Preiser, S. (1974). Regression und Recreation – Ein Beitrag zum Problem Stress und Erholung. *Zeitschrift fuer experimentelle und angewandte Psychologie*, 21, 575 – 591.

220. Loftus, G. R. (1993). A picture is worth a thousand p values: On the irrelevance of hypothesis testing in the microcomputer age. *Behavior Research Methods, Instruments & Computers*, 25, 250-256.

221. London, H., Schubert, D.S.P., & Washburn, D. (1972). Increase in autonomic arousal by boredom. *Journal of Abnormal Psychology*, 80, 29-36.

222. Lorist, M. M, Klein, M., Nieuwenhuis, S., De Jong, R., Mulder, G. & Meijman, T. F. (2000). Mental fatigue and task control: Planning and preparation. *Psychophysiology*, 37, 614-625.

223. Luchins, A. (1942). Mechanization in Problem Solving: The Effect of Einstellung. *Psychological Monographs*, 54(6), 1-94.

224. Luczak, H. (1975). *Untersuchungen informatorischer Belastung und* Beanspruchung [Examination of load and strain]. Forschungsbericht VDI-Z, Reihe 10, Nr. 2, Duesseldorf: VDI-Verlag.

225. Luczak, H. (1998). *Arbeitswissenschaft* [Ergonomics]. (2nd ed.). Berlin: Springer-Verlag.

226. Luczak, H. & Goebel, M. (2000). Signal Processing and Analysis in Application. In R. W. Backs & W. Boucsein (Eds.), *Engineering psychophysiology. Issues & Applications* (pp. 79-110). London: Lawrence Erlbaum Associates.

227. Luna, T. D. (1997). Air Traffic Controller Shiftwork: What are the implications for Aviation Safety? A Review. *Aviation, Space, and Environmental Medicine*, 68 (1), 69-79.

228. Luna, T. D., French, J., & Mitch, J. (1997). A Study of USAF Air Traffic Controller Shiftwork: Sleep, Fatigue, Activity and Mood Analyses. *Aviation, Space, and Environmental Medicine*, 68 (1), 18-23.

229. Lundberg, U. & Johansson, G. (2000). Stress and Health Risks in Repetitive Work and Supervisory Monitoring Work. In R. W. Backs & W. Boucsein (Eds.), *Engineering psychophysiology. Issues & Applications* (pp. 339- 359). London: Lawrence Erlbaum Associates.

230. Lykken, D. T., Rose, R., Luther, B., & Maley, M. (1966). Correcting psychophysiological measurements for individual differences in range. *Psychological Bulletin*, 66, 481-484.

231. Macdonald, W. & Bendak, S. (2000). Effects of workload level and 8- versus 12-h workday duration on test battery performance. *International Journal of Industrial Ergonomics*, 26, 399-416.

232. Mackie, R. R. (1987). Vigilance research – are we ready for countermeasures? *Human Factors*, 29, 707-723.

233. Mackworth, N. H. (1948). The breakdown of vigilance during prolonged visual search. *Quarterly Journal of Experimental Psychology*, 1, 6-21.

234. Makeig, S. & Inlow, M. (1993). Lapses in alertness: coherence of fluctuations in performance and the EEG spectrum. *Electroencephalography and Clinical Neurophysiology*, 86, 23-35.

235. Manning, C. A. (2000). *Measuring Air Traffic controller performance in a High-Fidelity Simulation* (DOT/FAA/AM-00/2). Washington, DC: FAA Office of Aviation Medicine.

236. Manning, C. A., Mills, S. H., Fox, C., Pfleiderer, E. M., & Mogilka, H. J. (2002). *Using Air Traffic Control Taskload Measures and Communication Events to Predict Subjective Workload* (DOT/FAA/AM-02/4). Washington, DC: Department of Transportation/Federal Aviation Administration, Office of Aviation Medicine.

237. Manning, C. & Stein, E. (2005). Measuring Air Traffic Controller Performance in the 21st Century. In B. Kirwan, M. Rodger, & D. Schaefer (Eds.) *Human Factors Impacts in Air Traffic Management* (pp. 283-316). Aldershot: Ashgate.

238. Manzey, D. (1998). Psychophysiologie mentaler Beanspruchung [Psychophysiology of mental strain]. In F. Roesler (Ed.), Ergebnisse und Anwendung der Psychophysiologie (Enzyklopaedie der Psychologie). Themenbereich C, Theorie und Forschung. Ergebnisse und Anwendungen der Psychophysiologie (Vol. 5, pp. 799-864). Goettingen: Hogrefe.

239. Masalonis, A. J., Callaham, M. B., & Wanke, C.R. (2003). *Dynamic Density and Complexity Metrics for Realtime Traffic Flow Management*. Presented at the ATM 2003 Conference, 5th EUROCONTROL/FAA ATM R&D Seminar (Budapest, Hungary).

240. Mascord, D. J. & Heath, R. A. (1992). Behavioral and physiological indices of fatigue in a visual tracking task. *Journal of Safety Research*, 23, 19-25.

241. Matthews, G. (2001). Levels of Transaction: A Cognitive Science Framework for Operator Stress. In: Hancock, P. (Eds.). *Stress, workload and fatigue* (pp. 5-33). Mahwah: Erlbaum.

242. Matthews, G., Jones, D., & Chamberlain, A. G. (1990). Refining the measurement of mood: The UWIST Mood Adjective Checklist. *British Journal of Psychology*, 81, 17-42.

243. Matthews, G., Joyner, L., Gilliland, K., Campbell, S., Huggins, J., & Falconer, S. (1999). Validation of a comprehensive stress state questionnaire: Towards a state „big three“? In I. Mervielde, I. J. Deary, F., DeFruyt, & F. Ostendorf (Eds.), Personality psychology in Europe (v.7), (pp.335-350). Tilburg: Tilburg University Press.

244. Matthews, G. & Zeidner, M. (2004). Traits, States, and the Trilogy of Mind: An Adaptive Perspective on Intellectual Functioning. In D. Y. Dai & R. J. Sternberg (Eds.) Motivation, Emotion, and Cognition (pp. 143-144). Mahwah,NJ: Lawrence Erlbaum.

245. Mavjee, V. & Horne, J. A. (1994). Boredom effects on sleepiness/alertness in the early afternoon vs. early evening and interactions with warm ambient temperature. *British Journal of Psychology*, 85, 317-333.

246. McAuley, E. (1994). Physical activity and psychosocial outcomes. In C. Bouchard, R. J. Shephard, & T. Stephens (Eds). *Physical activity, fitness, and health*. International proceedings and consensus statement. Champaign, IL : Human Kinetics, 551-568.

247. McBain, W. N. (1961). Noise, the arousal hypothesis, and monotonous work. *Journal of Applied Psychology*, 45, 309-317.

248. McBain, W. N. (1970). Arousal, monotony and accidents in line driving. *Journal of Applied Psychology*, 54,405-407.

249. McDonald, N. (1989). Fatigue and Driving. *Alcohol, Drugs and Driving*, 5, 185-191.

250. McDowell, R. J. S. & Wells, H. M. (1927). The physiology of monotony. *British Medical Journal*, 1, 414. In Barmack, 1939b.

251. McGrath, J. E. (1970). *Social and psychological factors in stress*. New York: Holt, Rinehart & Winston.

252. McGregor, D. K. & Stern, J. A. (1996). Time on task and blinks: effects on saccade duration. *Ergonomics*, 39, 649-660.

253. McKenna, F. P. (2002). Subjective measures: not perfect but what is? *Ergonomics*, 45(14), 998-1000.

254. McLean, G. A., Smith, L. T., Hill, T. J., & Rubenstein, C.J. (1993). *Physiological Correlates of Stress-Induced Decrements in Human Perceptual Performance.* (DOT/FAA/AM-93/19). Washington, DC: Department of Transportation/Federal Aviation Administration, Office of Aviation Medicine.

255. McEwen, B. S. & Wingfield, J. C. (2003). The concept of allostasis in biology and biomedicine. *Hormones and Behavior*, 43, 2-15.

256. Meier, F. (1984). Antriebsregulation und emotionale Beanspruchungseffekte bei gleichfoermiger kognitiver Arbeits-belastung [Strain effects of uniform cognitive task load]. *Zeitschrift fuer Arbeits- und Organisationspsychologie*, 28(4), 155-163.

257. Meijman, T. F. & Mulder, G. (1998). Mental workload. In P. J. Drenth, H. Thierry & C. de Wolff (Eds). *Handbook of work and organizational psychology, work psychology* (pp. 5-34). London: Psychology Press.

258. Melamed, S., Ben-Avi, I., Luz, J., & Green, M. S. (1995a). Objective and subjective work monotony: Effects on Job satisfaction, Psychological distress and absenteeism in blue-collar workers. *Journal of Applied Psychology*, 80 (1), 29-42.Melamed, S., Ben-Avi, I., Luz, J., & Green, M. S. (1995b). Repetitive Work, Work Underload and Coronary Heart Disease Risk Factors Among Blue-Collar Workers - The Cordis Study. *Journal of Psychosomatic Research*, 39(1), 19-29.

259. Melin, B., Lundberg, U., Derlund, J. S., & Granqvist, M. (1999). Psychological and physiological stress reactions of male and female assembly workers: a comparison between two different forms of work organization. *Journal of Organizational Behavior*, 20, 47-61.

260. Melton, C. E. (1982). *Physiological stress in air traffic controllers: A review* (FAA-AM-82-17). Washington, DC: Department of Transportation/Federal Aviation Administration, Office of Aviation Medicine.

261. Melton C. E. & Bartanowicz, R. S. (1986). Biological rhythms and rotating shift work: Some considerations for air traffic controllers and managers (DOT/FAA/AM-86/2). Washington, DC: U.S. Department of Transportation, Federal Aviation Administration.

262. Melton, C. E., McKenzie, J. M., Polis, B. D., Funkhouser, G. E. & Iampietro, P. F. (1971). *Physiological Responses in air traffic control personnel: O'Hare Tower.* (FAA-AM-71/2). Washington, DC: Department of Transportation/Federal Aviation Administration, Office of Aviation Medicine.

263. Melton, C. E., Smith, R. C., McKenzie, J. M., Wicks, S. M. & Saldiver, J. T. (1977). *Stress in Air Traffic Personnel: Low-Density towers and flight service stations* (FAA-AM-77/23). Washington, DC: Department of Transportation/Federal Aviation Administration, Office of Aviation Medicine.

264. Metzger, U. & Parasuraman, R. (2001). The Role of the Air Traffic Controller in Future Air Traffic Management: An empirical study of active control versus passive Monitoring. *Human Factors*, 43(4), 519-528.

265. Michielsen, H. J., De Vries, J., Van Heck, G. L., Van de Vijver, F. J. R., & Sijtsma, K. (2004). Examination of the Dimensionality of Fatigue: The Construction of the Fatigue Assessment Scale (FAS). *European Journal of Psychological Assessment*, 20(1), 39-48.

266. Mikulas, W. L. & Vodanovich, S. J. (1993). The essence of boredom. *Psychological Record*, 43(1), 3-12.

267. Mikulka, P. J., Scerbo, M. W., & Freeman, F. G. (2002). Effects of a Biocybernetic System on Vigilance Performance. *Human Factors*, 44(4), 654-664.

268. Ming, E. E., Adler, G. K., Kessler, R. C., Fogg, L. F., Matthews, K. A., Herd, J. A., & Rose, R. M. (2004). Cardiovascular Reactivity to Work Stress Predicts Subsequent Onset of Hypertension: The Air Traffic Controller Health Change Study. *Psychosomatic Medicine*, 66, 459-465.

269. Miyake, S. (2001). Multivariate workload evaluation combining physiological and subjective measures. *International Journal of Psychophysiology*, 40(3), 233-238.

270. Mogford, R. H., Guttman, J. A., Morrow, S. L., & Kopardekar, P. (1995). *The complexity Construct in Air Traffic Control: A review and synthesis of the Literature* (DOT/FAA/CT-TN95/22). Washington, DC: Department of Transportation/Federal Aviation Administration, Office of Aviation Research.

271. Mollard, R., Cabon, F., Bourgeois-Bougrine, S., & David., H. (2000). Measurement o f fatigue and sleepiness in ATC simulation. *Proceedings of the IEA 2000/HFES 2000 Congress* (Vol. 3, pp. 208-211), Santa Monica, CA: Human Factors and Ergonomics Society.

272. Monk, T. H. & Leng, V. C. (1982). Time of day effects in simple repetitive tasks. *Acta Psychologica*, 51, 207-221.

273. Moroney, W. F., Biers, D. W., & Eggemeier, F. T., 1995. Some measurement and methodological considerations in the application of subjective workload measurement techniques. *The International Journal of Aviation Psychology*, 5, 87-106.

274. Morris, T. L. & Miller, J. C. (1996). Electrooculographic and performance indices of fatigue during simulated flight. *Biological Psychology*, 42, 343-360.

275. Mulder, G. (1986). The concept of mental effort. In G. R. J. Hockey & A. W. K. Gaillard, & G. Mulder (Eds.), *Energetic and human information processing* (pp. 175-198). Dordrecht, The Netherlands: Martinus Nijhoff.

276. Mulder, G., Mulder, L. J., Meijman, T. F., Veldman, J. B. & Van Roon, A. M. (2000). A Psychophysiological Approach to Working Conditions. In R. W. Backs & W. Boucsein (Eds.), *Engineering psychophysiology. Issues & Applications* (pp. 139 - 160). London: Lawrence Erlbaum Associates.

277. Mulder, L. J. M., Van Roon, A., Veldman, H., Laumann, K., Burov, O., Quispel, L., & Hoogeboom, P. (2003). How to Use Cardiovascular state Changes in Adaptive Automation. In G. R. J. Hockey, A. W. K. Gaillard, & O. Burov (Eds.). *Operator Functional State* (pp. 260-272). Amsterdam: IOS Press.

278. Muse, L. A., Harris S. G., & Field H. S. (2003). Has the Inverted-U Theory of Stress and Job Performance Had a Fair Test? *Human Performance, 16*(4), 349-364.

279. Muzet, A. & Roge, J. (2003). Detecting Operator Drowsiness during a complex task through behavioural and physiological parameters. In G. R. J. Hockey, A. W. K. Gaillard, & O. Burov (Eds.). *Operator Functional State* (pp. 81-89). Amsterdam: IOS Press.

280. Nachreiner, F. (2002). UEber einige aktuelle Probleme der Erfassung, Messung und Beurteilung der psychischen Belastung und Beanspruchung [On some problems of assessing, measuring and evaluating mental workload]. *Zeitschrift fuer Arbeitswissenschaft*, 1-2, 10-21.

281. Nachreiner, F. & Schultetus, W. (2002). Normung im Bereich der psychischen Belastung – die Normen der Reihe DIN EN ISO 10075. *DIN-Mitteilungen*, 81.2002, 8, 519-533.

282. Neri, D. F., Oyung, R. L., Colletti, L. M., Mallis, M. M., Tam, P. Y., & Dinges, D. F. (2002). Controlled breaks as a fatigue countermeasure on the flight deck. A*viation, Space, and Environmental Medicine*, 3(7), 654–664.

283. Nickel, P., & Nachreiner, F. (2003). Sensitivity and diagnosticity of the 0.1 Hz component of heart rate variability as an indicator of mental workload. *Human Factors*, 45, 575–590.

284. Nickel, P., Nachreiner, F., Zdobych, A. & Yanagibori, R. (1998). Zur Beurteilung der psychischen Beanspruchung mit Hilfe der 0,1 Hz-Komponente der Herzfrequenzvariabilitaet – einige methodische und analysetechnische Probleme. *Zeitschrift fuer Arbeits- und Organisationspsychologie*, 42(4), 205 - 212.

285. Nunes, A. & Mogford, R. H. (2003). Identifying controllers strategies that support the ‚picture'. *Proceedings of the 47th Annual Meeting of the Human Factors and Ergonomics Society*. Santa Monica, CA: Human Factors and Ergonomics Society.

286. Oboznov, A. A., Yegorov, S. V., & Kostritsa, V. G. (1991). Mental image and the operator's reliability in monotony. *Soviet Journal of Psychology*, 12(2), 50-57.

287. O'Donnell, R. D. & Eggemeier, F. T. (1986). Workload assessment methodology. In K. R. Boff, L. Kaufman & J. P. Thomas (Eds.), *Handbook of perception and human performance.* Volume II, Cognitive processes and performance (pp. 42/1-42/49). New York: Wiley.

288. Oehman, A. (1988). Preattentive processes in the generation of emotions. In V. Hamilton, G. H. Bower & N. H. Frijda (Eds.), *Cognitive perspectives on emotion and motivation* (pp. 127–144). Dordrecht, The Netherlands: Martinus Nijhoff Publishers.

289. Oesterreich, R. (1999). Konzepte zu Arbeitsbedingungen und Gesundheit – Fuenf Erklaerungsmodele im Vergleich. In R. Oesterreich & W. Volpert (Eds). *Psychologie gesundheitsgerechter Arbeitsbedingungen (pp. 141-215).* Bern: Hans Huber Verlag.

290. Ognianova, V. M., Dalbokova, D. L., & Stanchev, V. (1998). Stress states, alertness and individual differences under 12-hour shiftwork. *International Journal of Industrial Ergonomics,* 21, 283-291.

291. O'Hanlon, J. F. (1981). Boredom: Practical consequences and a theory. *Acta Psychologica*, 49, 53-82.

292. Ohsuga, M., Shimono, F. & Genno, H. (2001). Assessment of phasic work stress using autonomic indices. *International Journal of Psychophysiology,* 40, 211-220.

293. Oweis, P. & Spinks, W. (2001). Biopsychological, affective and cognitive responses to acute physical activity. *Journal of sports medicine and physical fitness,* 41(4), 528-538.

294. Parasuraman, R. M., Molloy, R., & Singh, I. L. (1993). Performance consequences of automation induced "complacency". *International Journal of Aviation Psychology*, 3, 1-23.

295. Pennekamp, P., Bosel, R., Mecklinger, A. & Orr, H. (1994). Differences in EEG-theta for responded and omitted targets in a sustained attention task. *Journal of Psychophysiology*, 8, 131-141.

296. Perkins, R. E. & Hill, A. B. (1985). Cognitive and affective aspects of boredom. *British Journal of Psychology*, 76, 221-234.

297. Perneger, T. V. (1999). What's wrong with Bonferroni adjustments. *BMJ.* 316:1236-1238.

298. Persson, A., Wanek, B. & Johansson, A. (2001). Passive versus active operator work in automated process control - a job design case study in a control centre. *Applied Ergonomics*, 32, 441-451.

299. Petty, R. E. ,Fabrigar, L. R., Wegener, D. T., & Priester, J. R. (1996). Understanding Data When Interactions Are Present Or Hypothesized. *Psychological Science*, 7(4), 247-252.

300. Pfendl, C. (1985). Der Einfluss der Monotonie in einer UEberwachungsaufgabe auf die Entdeckungs- und Aufmerksamkeitsleistung [The influence of monotony on detection and vigilance performance during a monitoring task]. *Zeitschrift fuer Arbeitswissenschaft*, 39(4), 231-236.

301. Pistre, M. & Bubalo, V. (2004). *Guidelines for ATCO Common Core Content Initial Training* (HRS/TSP-002-GUI-04). Brussels: EUROCONTROL.

302. Plath, H. E. & Richter, P. (1984). Der BMS-II-Erfassungsbogen – ein Verfahren zur skalierten Erfassung arbeitsbedingter Beanspruchungsfolgen bei UEberwachungs- und Steuertaetigkeiten [BMSII - A scale for the assessment of work-related strain in monitoring and controlling tasks]. *Psychologie fuer die Praxis,* Ergaenzungsband, 64–70.

303. Plutchik, R. (1980). *Theories of emotion*. New York: Academic.

304. Porterfield, D. H. (1997). Evaluating controller communication time as a measure of workload. *International Journal of Aviation Psychology*, 7, 171-182

305. Posner, M. I. & Rafal, R. D. (1987). Cognitive theories of attention and the rehabilitation of attentional deficits. In R. J. Meier, A. L. Benton & L. Diller (Eds.). *Neuropsychological rehabilitation* (pp. 182-201). Edinburgh: Churchill Livingstone.

306. Pribram, K., & McGuinnes, D. (1975). Arousal, activation, and effort in the control of attention. *Psychological Review,* 159, 82(2) 116-149.

307. Prinzel, L. J, Sawin, D. A., & Scerbo, M. W. (1995). *Boredom in vigilance: Stuck in a rut.* Paper presented at the second annual Mid-Atlantic Human Factors Conference. Reston, VA.

308. Rau, R. & Richter, P. (1996). Psychophysiological analysis of strain in real life work situations. In J. Fahrenberg & M. Myrtek (Eds.). *Ambulatory Assessment* (pp. 271-285). Seattle: Hogrefe & Huber Publishers.

309. Reason, J. (1992). Cognitive underspecification: Its variety and consequences. In B. J. Baars (Ed.) *Experimental Slips and Human Error: Exploring the Architecture of Volition* (pp. 71-91). New York: Plenum Press.

310. Redding, R. (1992). Analysis of operational errors in Air traffic control. *Proceedings of the Human Factors Society Annual Meeting*, Vol. 2 (pp. 1321-1325). Santa Monica, CA: Human Factors Society.

311. Reid, G. B., & Nygren, T. E. (1988). The Subjective Workload Assessment Technique: A Scaling Procedure for Measuring Mental Workload. In P. Hancock and N. Meshkati (Eds.), *Human Mental Workload* (pp. 185-218). Amsterdam, the Netherlands: North Holland.

312. Revers, W.J. (1949). *Die Psychologie der Langeweile [The Psychology of Boredom]*. Meisenhein: Hain-Verlag.

313. Richter, P., Debitz, U. & Schulze, F. (2002). Diagnostik von Arbeitsanforderungen und kumulativen Beanspruch-ungsfolgen am Beispiel eines Call-Centers [Diagnostic of job demands and cumulating strain consequences in call center jobs]. *Zeitschrift fuer Arbeitswissenschaft*, 1-2, 67- 75

314. Richter, P. & Hacker, W. (1998). *Belastung und Beanspruchung: Stress, Ermuedung und Burnout im Arbeitsleben.* [Strain and load: stress, fatigue and burnout at work]. Heidelberg: Asanger.

315. Riley, V. & Parasuraman, R. (1997). Humans and Automation: Use, Misuse, Disuse, Abuse. *Human Factors*, 39(2), 230-253.

316. Roach, G. D., Reid, K. J., & Dawson, D. (2003). The amount of sleep obtained by locomotive engineers: effects of break duration and time of break onset. *Occupational and Environmental Medicine*, 60, e17. Retrieved November 8, 2005, from http://oem.bmjjournals.com/.

317. Robbins, T. W. (1997). Arousal systems and attentional processes. *Biological Psychology*, 45, 57-71.

318. Robertson, I. H., Manly, T., Andrade, J., Baddeley, B. T., & Yiend, J. (1997). 'Oops!': performance correlates of everyday attentional failures in traumatic brain injured and normal subjects. *Neuropsychologia*, 35, 747-758.

319. Rockstuhl, T. (2002). *BMS-II- Modifikationen und Beanspruchungsmessung in Call Centers [Modifications of BMS-II and the measurement of strain in call centers]*. Unpublished thesis, University of Dresden.

320. Rodgers, M. D., Mogford, R. H. & Mogford, L. S. (1998*). The Relationship of Sector Characteristics to Operational Errors* (DOT/FAA/AM-98/14). Washington, DC: Department of Transportation/Federal Aviation Administration, Office of Aviation Medicine.

321. Rogé, J., Pebayle, T., & Muzet, A. (2001). Variations of the level of vigilance and of behavioural activities during simulated automobile driving. *Accident Analysis & Prevention*, 33(2), 181-186.

322. Rohmert, W. (1973). Problems of determination of rest allowances. Part 1: Use of modern methods to evaluate stress and strain in static muscular work. *Applied Ergonomics*, 4(2),91-95.

323. Rohmert, W. (1984). Das Belastungs-Beanspruchungs-Konzept [The stress-strain-concept]. *Zeitschrift fuer Arbeitswissenschaft*, 4, 193-200.

324. Rohmert, W. & Luczak, H. (1973). Ergonomische Beurteilung informatorischer Arbeit [Ergonomic evaluation of cognitive tasks]. *Internationale Zeitschrift fuer angewandte Physiologie*, 31, 209 -229.

325. Rose, C. L., Murphy, L. B., Byard, L. & Nikzad, K. (2002). The role of the Big Five personality factors in vigilance performance and workload. *European Journal of Personality*, 16(3), 185-200.

326. Rose, R. M. & Fogg, L . F. (1993). Definition of a responder: analysis of behavioral, cardiovascular, and endocrine responses to varied workload in air traffic controllers. *Psychosomatic Medicine*, 55(4), 325-338.

327. Rosekind, M. R., Gander, P. H., Gregory, K. B. et al. (1996a). Managing fatigue in operational settings 1: Physiological considerations and countermeasures. *Behavioral Medicine*, 21, 157-165.

328. Rosekind, M. R., Gander, P. H., Gregory, K. B. et al. (1996b). Managing fatigue in operational settings 2: An integrated approach. *Behavioral Medicine*, 21, 166-170.

329. Rosekind, M. R., Gander, P. H., Miller, D. L., Gregory, K. B., Smith, R. M., Weldon, K. J., Co, E. L., McNally, K. L., & Lebacqe, J. V. (1994). Fatigue Operational Settings: Examples from the Aviation Environment. *Human Factors*, 36(2), 327-338.

330. Rosnow, R. L. & Rosenthal, R. (1989). Statistical procedures and the justification of knowledge in psychological science. *American Psychologist*, *44*, 1276-1284.

331. Russo, M., Thomas, M., Thorne, D., Sing, H., & Redmond, D. (2003). Oculomotor impairment during chronic partial sleep deprivation. *Clinical Neurophysiology*, 114, 723-736.

332. Ryan, T. A. (1947). *Work and Effort.* New York.

333. Rzepa, T. (1984). Typological determinants of operator functioning in monotonous work conditions. *Polish Psychological Bulletin*, 15(2), 135-141.

334. Salvendy, G. (2002). Use of subjective rating scores in ergonomics research and practice. *Ergonomics*, 45(14), 1005-1007.

335. Sammer, G. (1998). Heart rate variability and respiratory changes associated with physical and mental load: a non linear analysis. *Ergonomics*, 41(5), 746-755.

336. Sanders, A. F. (1983). Towards a model of stress and human performance. *Acta Psychologica*, 53, 61-97.

337. Sanders, A. F. (1997). A summary of resource theories from a behavioural perspective. *Biological Psychology*, 45, 5-18.

338. Sawin, D. A. & Scerbo, M. W. (1995). Effects of instruction type and boredom proneness in vigilance: Implications for boredom and workload. *Human Factors*, 37(4), 752-765.

339. Scerbo, M. W. (2001). Stress, workload and boredom in vigilance: a problem and an answer. In P. A. Hancock & P. Desmond (Eds.). *Stress, workload and fatigue* (pp. 267-278). Mahwah: Erlbaum.

340. Scerbo, M. W., Greenwald, C. Q., & Sawin, D. A. (1993). The effects of subject-controlled pacing and task type upon sustained attention and subjective workload. *Journal of General Psychology*, 113, 293-307.

341. Scerbo, M. W., Rettig, K. M., & Bubb-Lewis, C. L. (1994). A validation study of a task-related boredom scale. In *Proceedings of the 2nd Mid-Atlantic Human Factors Conference* (pp. 135-136). Washington, D. C.

342. Scerbo, M. W. & Sawin, D. A. (1994). *Vigilance: How much is enough?* Paper presented at the second annual Mid-Atlantic Human Factors Conference. Reston, VA.

343. Schaefer, D., Flynn, M. & Skraaning, G. (2003). The Presentation of Conflict Resolution Advisories to Air Traffic Controllers – a Human Factors Perspective. In D. Harris, V. Duffy, M. Smith, & Stephanidis, C. (Eds.) *Human-Centred computing. Cognitive, social and ergonomic Aspects*, Vol. 3 of the proceedings of HCI 2003, 10th International Conference of Human-Computer Interaction, 5th International Conference on Engineering Psychology (pp. 572-576). Mahwah: Lawrence Erlbaum Associates.

344. Schaefer, D. & Smith, D. (2006). Human-in-the-Loop Studies in the EUROCONTROL Human Factors Laboratory. *Transactions of The Society for Modeling and Simulation International*, Special Issue: Simulation of Air Traffic (in press).

345. Schapkin, S. A. & Gusev, A. N. (2003). Operator Functional State and Vigilance: Mediating Effect of Brain Hemispheres. In G. R. J. Hockey, A. W. K. Gaillard & O. Burov (Eds.). *Operator Functional State* (pp.140-150). Amsterdam: IOS Press.

346. Schellekens, J. M. H., Sijtsma, G. J., Vegter, E., & Meijman T. F. (2000). Immediate and delayed after-effects of long lasting mentally demanding work. *Biological Psychology*, 53(1), 37-56.

347. Scheuch, K. (1986). Psychophysiologie und Arbeitsmedizin [Psychophysiology and occupational medicine]. *Zeitschrift fuer Psychologie*, 194, 343-353.

348. Schmidtke, H. (2002). Vom Sinn und Unsinn der Messung psychischer Belastung und Beanspruchung [Sense and Nonsense of the measurement of mental strain and workload]. Zeitschrift fuer Arbeitswissenschaft, 1-2, 4-9.

349. Schoenpflug, W. (1986). Effort regulation and individual differences in effort expenditure. In G. R. J. Hockey, A. W. K. Gaillard, & M. G. H. Coles (1986). *Energetics and human information processing* (pp.271-284). Dordrecht: Matinus Nijhoff.

350. Schoenpflug, W. (1987). Beanspruchung und Belastung bei der Arbeit. Konzepte und Theorien. In *Enzyklopaedie der Psychologie, Themenbereich D, Serie III, Arbeitspsychologie (*Vol 1, pp. 130-84). Goettingen: Hogrefe.

351. Schroeder, D. J. (1982). *The loss of prescribed separation between aircraft: how does it occur?* Proceedings (P-114) Behavioral objectives in Aviation Automated Systems Symposium (pp. 257-269). Washington DC, Society for Automotive Engineers.

352. Schroeder, D. J., Touchstone, R. M., Stern, J. A., Stoliarov, N., & Thackray, R. (1994). *Maintaining vigilance on a simulated ATC monitoring task across repeated sessions (*DOT/FAA/AM-94/6). Office of Aviation Medicine, Federal Aviation Administration, Washington, D.C.

353. Schuette, M. (1999). Mental strain and the problem of repeated measurements. *Ergonomics*, 42(12), 1665-1678.

354. Schuette, M. (2002). Bestimmung der bedingungsbezogenen Messgenauigkeit der Anstrengungsskala. [Deter-mination of the condition related reliability of the effort scale]. *Zeitschrift fuer Arbeitswissenschaft*, 1-2, 37-45.

355. Schuhfried, G. & Prieler, J. (2000). *Vienna Reaction Test – Test manual*. Moedling: Schuhfried.

356. Schultz-Hardt, S., Meinken, I., Rott, A., & Frey, D. (2001). Psychische Saettigung: Eine neue experimentelle Untersuchung zu einem alten Konstrukt [Mental satiation: a new experimental study of an old construct]. *Zeitschrift fuer Experimentelle Psychologie*, 48(3), 188-200.

357. Schultz-Hardt, S., Rott, A., Meinken, I., & Frey, D. (2001). Ein weiterentwickeltes Modell psychischer Saettigung [A further developed model of satiation]. *Psychologische Rundschau*, 52(3), 141-149.

358. Sedlmeier, P. (1996). Jenseits des Signifikanztest-Rituals: Ergaenzungen und Alternativen [Beyond the Ritual of Significance Testing: Alternative and Supplementary Methods]. *Methods of Psychological Research Online*, 1(4), 41-63. Retrieved July, 3, 2004, from http://www.pabst-publishers.de/mpr/.

359. Seibt, R., Boucsein, W., & Scheuch, K. (1998). Effects of different stress settings on cardiovascular parameters and their relationship to daily life blood pressure in normotensives, borderline hypertensives and hypertensives. *Ergonomics*, 41(5), 634-648.

360. Shirom, A., Westman, M., & Melamed, S. (*1999).* The Effects of Pay Systems on Blue-Collar Employees' Emotional Distress: The Mediating Effects of Objective and Subjective Work Monotony. *Human Relations, 52(8), 1077-1097.*

361. Sirevaag, E. J. & Stern, J. A. (2000). Ocular Measures of Fatigue and Cognitive Factors. In R. W. Backs & W. Boucsein (Eds.), *Engineering psychophysiology. Issues & Applications* (pp. 269 - 287). London: Lawrence Erlbaum Associates.

362. Skraaning Jr., G. (2003). *Experimental Control versus Realism: Methodological Solutions for Simulator Studies in Complex Operating Environments*. Unpublished doctoral dissertation. ECD Halden Reactor Project, University of Science and Technology, Norway.

363. Sluiter, J. K., De Croon, E. M., Meijman, T. F., & Frings-Dresen, M. H. W. (2003). Need for recovery from work related fatigue and its role in the development and prediction of subjective health complaints. *Occupational and Environmental Medicine*, 60, 62-70.

364. Sluiter, J. K., Frings-Dresen, M. H., Meijman, T. F., & Van der Beek, A. J. (2000). Reactivity and recovery from different types of work measured by catecholamines and cortisol: a systematic literature overview. *Occupational and Environmental Medicine*, 57(5), 298-315.

365. Smit, A. S., Eling, P. A. T. M., Coenen, A. M. L. (2004a). Mental effort causes vigilance decrease due to resource depletion. *Acta Psychologica*, 115, 35-42.

366. Smith, P. C. (1955). The prediction of individual differences in susceptibility to industrial monotony. *Journal of Applied Psychology*, 39, 322-329.

367. Smith, R. C. (1980). *Stress, Anxiety and the Air Traffic Control Specialist: Some Conclusions from a Decade of Research.* (FAA-AM-80-14). Washington, DC: Department of Transportation/Federal Aviation Administration, Office of Aviation Medicine.

368. Smith, R.P. (1981). Boredom. A Review. *Human Factors*, 23(3), 329-340.

369. Smolensky, M. W. (1990). The effect of work load history on operational errors in air traffic control simulation: The hysteresis effect--expectancy perseverance or short-term memory overload? Texas Tech U, US.

370. Soames-Job, R. F. & Dalziel, J. (2001). Defining fatigue as a condition of the organism and distinguishing it from habituation, adaptation and boredom. In P. A. Hancock & P. Desmond (Eds.). *Stress, workload and fatigue* (pp. 466-478). Mahwah: Erlbaum.

371. Spencer, M. B., Rogers, A. S., & Stone, B. M. (1997). *A Review of the Current Scheme for the Regulation of Air Traffic Controllers Hours (SCRATCOH)*. Farnborough, UK: Defense Evaluation and Research Agency PLSD/CHS5/CR/97/020.

372. Sperandio, J. C. (1971). Variation of operator's strategies and regulating effects on workload. *Ergonomics*, 14(5), 571-577.

373. Sperandio, J. C. (1978). The regulation of working methods as a function of workload among air traffic controllers. *Ergonomics, 21*, 195-202.

374. Stachowiak, H. (1973). *Allgemeine Modelltheorie.* Wien: Springer.

375. Stager, P., Hameluck, D., & Jubis, R. (1989). *Underlying factors in air traffic control incidents.* In Proceedings of the Human Factors Society 33rd Annual Meeting (Vol. 1, pp. 43-46). Santa Monica, CA, USA: Human Factors Society.

376. Stemmler, G. (2001). Grundlagen psychophysiologischer Methodik. In F. Roesler (Ed.), *Enzyklopaedie der Psychologie,* Einfuehrung in die Psychophysiologie Themenbereich C, Theorie und Forschung, Serie 1, Biologische Psychologie, Bd. 4, Grundlagen und Methoden der Psychophysiologie (pp. 1-70). Goettingen: Hogrefe.

377. Stern, J. A. (1964). Towards a definition of psychophysiology. *Psychophysiology,* 1, 90-91.

378. Stern, J. A., Boyer, D., & Schroeder, D. J. (1994). *Blink Rate As a Measure of Fatigue: A Review.* (DOT/FAA/AM 94/17). Washington, DC: Department of Transportation/Federal Aviation Administration, Office of Aviation Medicine.

379. Stern, J. A., Boyer, D., Schroeder, D. J., Touchstone, R. M., & Stoliarov, N. (1994). *Blinks, Saccades, and Fixation Pauses During Vigilance Task Performance: I. Time on Task* (FAA-AM-94/26). Washington, DC: Department of Transportation/Federal Aviation Administration, Office of Aviation Medicine.

380. Stern, J. A., Boyer, D., Schroeder, D. J., Touchstone, R. M., & Stoliarov, N. (1996*). Blinks, Saccades, and Fixation Pauses During Vigilance Task Performance: II. Gender and Time of Day* (FAA-AM-96/9). Washington, DC: Department of Transportation/Federal Aviation Administration, Office of Aviation Medicine.

381. Sternbach, R. A., Alexander, A. A., Rice, D. G., & Greenfield, N. S. (1969). Technical note: some views on the treatment of artifacts in psychophysiological recordings. *Psychophysiology*, 6(1), 1-5.

382. Stirnman, M., Rothmann, V., Graham, R., Dowdall, R., & R. Eveleigh (2005*). C-ATM High-Level Operational Concept. Deployable From 2012.* Version 1.2. CATM-WP111-ERC-HLOC-D111-V0120. EUROCONTROL.

383. Stokes, A. & Kite, K. (1994). *Flight Stress: Stress, Fatigue and Performance in Aviation.* Cambridge: Great Britain at the University Press.

384. Straussberger, S. & Kallus, W. (2003). Beanspruchungsoptimierung durch Pausengestaltung bei komplexen UEberwachungstaetigkeiten. [Workload optimization through structured pauses in complex air traffic control tasks] In P. Richter & K. Westhoff (Host-Eds.), *Wirtschaftspsychologie - Flexibilisierung der Arbeit,* (Vol. 1, pp. 157-160). Lengerich: Pabst

385. Strelau, J. (2001). The concept and status of traits in research on temperament. *European Journal of Personality*, 15, 311-325.

386. Swinkels, A. & Giuliano, T. A. (1995). The measurement and conceptualization of mood awareness: Monitoring and labeling one's mood states. *Personality and Social Psychology Bulletin, 21,* 934-949.

387. Szalma, J. L. & Hancock, P. A. (2002). *On mental resources and performance under stress*. Unpublished white paper, MIT2 Laboratory, University of Central Florida. Available at www.mit.ucf.edu.

388. Szalma, J. L., Warm, J. S., Matthews, G., Dember, W. N., Weiler, E. M., & Eggemeier, F. T. (2004). Effects of Sensory Modality and Task Duration on Performance, Workload, and Stress in Sustained Attention. *Human Factors*, 46(2), 218-233.

389. Taillard, J., Philip, P., Chastang J. F., & Bioulac, B. (2004). Validation of Horne and Ostberg morningness-eveningness questionnaire in a middle-aged population of French workers. *Journal of Biological Rhythms*, 19(1), 76-86.

390. Tanaka, Y. & Yamaoka, K. (1993). Blink activity and task difficulty. *Perceptual and Motor Skills,* 77, 55-66.

391. Task Force of The European Society of Cardiology and The North American Society of Pacing and Electrophysiology (1996). Heart rate variability. Standards of measurement, physiological interpretation, and clinical use. *European Heart Journal,* 17, 354-381.

392. Tattersall, A. J. & Hockey, G. R. J. (1990). The assessment of workload in a complex monitoring and fault diagnosis test. In D. Brogan (Ed.) *Visual search. Visual perception.* London: Taylor & Francis.

393. Tattersall, A.J. & Foord, P. S. (1996). An experimental evaluation of instantaneous self-assessment as a measure of workload. *Ergonomics,* 39(5), 740-748.

394. Thackray, R. I. (1981a). *Boredom and monotony as a consequence of automation: A consideration of the evidence relating boredom and monotony to stress.* (FAA-AM-81-1). Washington, DC: Department of Transportation/Federal Aviation Administration, Office of Aviation Medicine.

395. Thackray, R. I. (1981b). The stress of boredom and monotony: a consideration of the evidence. *Psychosomatic Medicine*, 43, 165-176.

396. Thackray, R. I., Bailey, J. P., & Touchstone, R. M. (1975). *Physiological, subjective, and performance correlates of reported boredom and monotony while performing a simulated radar control task.* (FAA-AM-75-8). Washington, DC: Department of Transportation/Federal Aviation Administration, Office of Aviation Medicine.

397. Thackray, R. I. & Touchstone, R. M. (1983). *Rate of initial recovery and subsequent radar monitoring performance following a simulated emergency involving startle* (FAA-AM-83-13). Washington, DC: Department of Transportation/Federal Aviation Administration, Office of Aviation Medicine.

398. Thayer, J. F. & Friedman, B. H. (2000). The Design and Analysis of Experiments in engineering psychophysiology. In R. W. Backs & W. Boucsein (Eds.), *Engineering psychophysiology. Issues & Applications* (pp. 59 - 78). London: Lawrence Erlbaum Associates.

399. Thayer, R. E. (1967). Measurement of activation through self-report. *Psychological Reports*, 20, 663-678.

400. Thayer, R. E. (1987). Energy, Tiredness, and Tension Effects of a sugar snack versus moderate exercise. *Journal of Personality and Social Psychology*, 52(1), 119-125.

401. Thayer, R. E., Newman, J. R. & McClain, T. M. (1994). Self-regulation of mood: Strategies for changing a bad mood, raising energy, and reducing tension. *Journal of Personality and Social Psychology*, 67, 910-925.

402. Thiffault, P. & Bergeron, J. (2003a). Fatigue and individual differences in monotonous simulated driving. *Personality and Individual Differences*, 34(1), 159-176.

403. Thiffault, P. & Bergeron, J. (2003b). Monotony of road environment and driver fatigue: a simulator study. *Accident Analysis and Prevention*, 35, 381-391.

404. Touzé, P. A. (2005). Personality and performance. *Le travail human*, 28(1), 38-53.

405. Tsang, P. S. & Vidulich, M. A. (2002). *Principle and Practice of Aviation.* Mahwah, NJ: Lawrence Erlbaum.

406. Udris, I. & Alioth, A. (1980). Fragebogen zur Subjektiven Arbeitsanalyse (SAA) [Questionnaire for Subjective Work Analysis]. In E. Martin, I. Udris, U. Ackermann, & K. Oegerli (Eds.) *Monotonie in der Industrie* (pp. 61-68; 204-207), Bern: Huber.

407. Ulich, E. (1960). Unterforderung als arbeitspsychologisches Problem [Underload as a problem of work psychology]. *Psychologie und Praxis*, 4, 156-161.

408. Ulich, E. (1998). *Arbeitspsychologie*. (4th ed.) Stuttgart: Schaeffer Poeschel.

409. Van der Flier, H. & Schoonman, W. (1988) Railway signals passed at danger: situational and personal factors underlying stop signal abuse. *Applied Ergonomics*, 19, 135-141.

410. Van der Hulst, M., Meijman, T., & Rothengatter, T. (2001). Maintaining task set under fatigue: a study of time-on-task effects in simulated driving. *Transportation Research, Part F: Traffic Psychology and Behaviour*, 4(2), 103-118.

411. Van der Linden, D., Frese, M., & Meijman, T. F. (2003a). Mental fatigue and the control of cognitive processes: effects on perseveration and planning. *Acta Psycologica*, 112(1), 45-65.

412. Van der Linden, D., Frese, M., & Sonnentag, S. (2003b). The impact of mental fatigue on exploration in a complex computer task: rigidity and loss of systematic strategies. *Human Factors*, 45(3), 483-494.

413. Van Dijk, F. J. H. & Swaen, G. M. H. (2003). Fatigue at work. *Occupational and Environmental Medicine*, 60(1), 1-2.

414. Van Orden, K. F, Limbert, W., Makeig, S. & Jung, T. (2001). Eye Activity Correlates of Workload during a Visuospatial Memory Task. *Human Factors*, 45(1), 111-121.

415. Van Veldhofen, M. & Broersen, S. (2003). Measurement quality and validity of the “need for recovery Scale”. *Occupational and Environmental Medicine*, 60(1), i3-9.

416. Van Zomeren, A. H. & Brouwer, W. H. (1994). *Clinical Neuropsychology of Attention.* New York: Oxford University Press.

417. Veltman, H. & Jansen, C. (2003). Differentiation of Mental Effort measures: Consequences for Adaptive Automation. In G. R. J. Hockey, A. W. K. Gaillard, & O. Burov (Eds.). *Operator Functional State* (pp. 249-259). Amsterdam: IOS Press.

418. Veltman, H. J. A. & Jansen, C. (2004). *The Adaptive Operator*. Presented at the Human Performance, Situation Awareness and Automation Technology Conference. March 22-25, 2004 Daytona Beach FL.

419. Veltman, J. A. & Gaillard, A. W. K. (1998). Physiological workload reactions to increasing levels of task difficulty. *Ergonomics*, 41(5), 656-669.

420. Veltman, J. A. & Gaillard, A. W. K. (1996). Physiological indices of workload in a simulated flight task. *Biological Psychology*, 42, 323-342.

421. Verwey, W. B. & Zaidel, D. M. (1999). Preventing drowsiness accidents by an alertness maintenance device. *Accident Analysis and Prevention*, 31, 199-211.

422. Vidulich, M. A. (2000). The relationship between mental workload and situation awareness. *Proceedings of the IEA 2000/HFES 2000 Congress Volume 3* (pp. 460-463). Santa Monica, CA: Human Factors and Ergonomics Society,.

423. Vodanovich, S. (2003). Psychometric Measures of Boredom. A Review. *The Journal of Psychology*, 137(6), 569-595.

424. Vodanovich, S.J., & Kass, S. J. (1990). A factor analytic study of the Boredom Proneness Scale. *Journal of Personality Assessment*, 55, 115-123.

425. Vodanovich, S. J., Verner, K. M., & Gilbride, T. V. (1991). Boredom proneness: Its relationship to positive and negative affect. *Psychological Reports*, 69, 1139-1146.

426. Voelk, Wolfgang. (1988). Aufgabenvariation unter Laerm - Ein Beitrag zur Untersuchung von Regelprozessen bei monotoner Taetigkeit [Variation of task under the influence of noise]. Dissertation, Universitaet Wuerzburg, Germany.

427. Vogt, J., Adolph, L., Azan, T., Udovic, A., & Kastner, M. (2002). Stress in modern air traffic control systems and potential influences on memory. *Human Factors and Aerospace Safety* 2(4), 355-378.

428. Vogt, J. & Leonhardt J. (2005). Increasing safety by stress management. *Safety Science Monitor*, 9(1), Art. 6.

429. Vogt, J., Leonhardt, J., Koper, B., & Pennig, S. (2004). Economic evaluation of CISM-a pilot study. *International Journal of Emergency Mental Health*, 6(4),185-196.

430. Vormayr, E., Kallus, K. W. & Hoffmann, P. (2005). Computerized version of the Safety-Efficiency-Taskload-Workload (SET-W) rating. Karl-Franzens-University of Graz.

431. Wagner, M. & Karner, T. (2003). Cognitrone - Test manual. Moedling: Schuhfried

432. Wagner, T., Rudolf, M., Noack, F. (1998). Die Herzfrequenzvariabilitaet in der arbeitspsychologischen Feldforschung – Methodenprobleme und Anwendungsbeispiele [Heart rate variability in applied research – probems related to methodology and some examples]. *Zeitschrift fuer Arbeits- und Organisationspsychologie*, 42 (4), 197-204.

433. Walker, C. R. & Guest, R. H. (1952). The man on the assembly line. *Harvard Business Review*, 30, 71-83.

434. Wallace, D., & Green, S. B. (2002). Analysis of repeated measures designs with linear mixed models. In D. S. Moskowitz & S. L. Hershberger (Eds.), *Modeling intraindividual variability with repeated measures data* (pp. 103-34). Mahwah, NJ: Erlbaum.

435. Wallace, J. C., Vodanovich, S. J., Restino, B. M. (2003). Predicting cognitive failures from boredom proneness and daytime sleepiness scores: an investigation within military and undergraduate samples. *Personality and Individual Differences*, 34, 635-644.

436. Walschburger, P. (1976). *Zur Beschreibung von Aktivierungsprozessen - Eine Methodenstudie zur psychophysiologischen Diagnostik [The description of activation processes]*. Unpublished Dissertation, Albert-Ludwig-Universitaet, Freiburg.

437. Warm, J. S. (1993). Vigilance and target detection. In B. M. Huey & C. D. Wickens (Eds.), *Workload transition: Implications for individual and team performance* (pp. 139-170). Washington, DC: National Academy.

438. Watt, J. D. (2002). Fighting more than fires: Boredom proneness, workload stress, and underemployment among urban firefighters. Dissertation-Abstracts-International:-Section-B:-The-Sciences-and-Engineering. 63(5-B): 2637.

439. Watt, J. D. & Blanchard, M. J. (1994). Boredom Proneness and the Need for Cognition. *Journal of Research in Personality*, 28(1), 44-51.

440. Weber, A., Jermini, C. & Grandjean, E. P. (1975). Relationship between objective and subjective assessment of experimentally induced fatigue. *Ergonomics*, 18, 151-156.

441. Weikert, C. & Johansson, C. R. (1999). Analyzing Incident Reports for Factors contributing to Air Traffic Control related Incidents. In *Proceedings of the Human Factors and Ergonomics Society 43rd Annual Meeting* (pp. 1075-1079). Santa Monica, CA: Human Factors and Ergonomics Society.

442. Welford, A. T. (1965). Fatigue and monotony. In O. G. Edholm & A. L. Bacharach (Eds.), *The Physiology of Human Survival*. London: Academic Press.

443. Whitfield, D. (1979). A preliminary study of the air traffic controller's picture. *Journal of the Canadian air traffic controllers' association*, 11(1) 19-22,25,28.

444. Whitfield, D. & Jackson, A. (1982). The air traffic controller picture as an example of the mental model. In G. Johannsen & J. E. Rijnsdorp (Eds.). *Proceedings of the IFAC Conference on Analysis, Design, and Evaluation of Man-Machine Systems* (pp. 45-52). London: Pergamon Press.

445. Wickens, C. D. & Hollands, J. G. (2000). *Engineering Psychology and Human Performance.* 3rd Edition, Upper Saddle River, NJ: Prentice.

446. Wieland-Eckelmann, R. (1997). Kognitive Grundfunktionen: Aufmerksamkeit [Cognitive Functions: Attention]. In Luczak & Volpert (Eds.). *Handbuch Arbeitswissenschaft* (pp. 431-434). Stuttgart: Schaeffer-Poeschel.

447. Wieland-Eckelmann, R. & Baggen, R. (1994). Beanspruchung und Erholung im Arbeits-Erholungs-Zyklus [Strain and recovery during work and in leisure]. In R. Wieland-Eckelmann (Ed.). *Erholungsforschung* (pp. 102-154). Weinheim: Beltz-Verlag.

448. Wilder, J. (1957). The Law of Initial Values in neurology and psychiatry: Facts and problems. *Journal. of Nervous and Mental Diseases*, 125, 73-86.

449. Willems, B. & Truitt, T. R. (1999). *Implications of Reduced Involvement in En Route Air Traffic Control.* DOT/FAA/CT-TN99/22.

450. Wilson, G. F. & Russell, C. A. (2003a). Operator Functional State Classification using multiple psychophysiological features in an Air Traffic Control Task. *Human Factors*, 45(3), 381-389.

451. Wilson, G. F. & Russell, C. A. (2003b). Real-time-assessment of mental workload using psychophysiological measures and artificial neural networks. *Human Factors*, 635-643.

452. Wundt, W. (1896). *Grundriss der Psychologie* [Basics of Psychology]. Leipzig: Wilhelm Engelmann.

453. Wyatt, S. (1929). Boredom in industry. *Personnel Journal*, 8, 161-171.

454. Wyatt, S., Fraser, J. A., & Stock, F. G. L., (1929). The effects of monotony and work. London: Medical Research Council, Industrial Fatigue Research Board, His Majesty's Stationery Office. In Davies et al.,1983.

455. Xing, J. & Bailey, L. (2005). ATC Operational Errors: Exceeding the Limits of Cognitive Capacities. *Proceedings of the 13th International Symposium on Aviation Psychology* (pp. 651-656). Dayton, OH: Wright State University.

456. Yamada, F. (1998). Frontal midline theta rhythm and eyeblinking activity during a VDT task and a video game: useful tools for psychophysiology in ergonomics. *Ergonomics*, 41(5), 678-688.

457. Yamamoto S., Matsuoka S., & Ishikawa T. (1989). Variations in EEG Activities During VDT Operation. Proceedings of the third international conference on human-computer interaction, Vol.1 on Work with computers: organizational, management, stress and health aspects (pp. 225-232), Boston, USA.

458. Yerkes, R.M. & Dodson J.D. (1908). *The Relation of Strength of Stimulus to Rapidity of Habit-Formation. Journal of Comparative Neurology and Psychology*, 18, 459-482

459. Young, M. & Clynick, G. (2005). A test flight for malleable attentional ressources theory. In P. D. Bust & P. T. McCabe (Eds). *Contemporary Ergonomics 2005* (pp. 548-552). Taylor & Francis: London.

460. Young, M. S. & Stanton, N. A. (2002a). Malleable Attentional Resources Theory: A new explanation for the effects of mental underload for performance. *Human Factors*, 44(3), 365-375.

461. Young, M. S. & Stanton, N. A. (2002b). Attention and automation: New perspectives on mental underload and performance. *Theoretical Issues in Ergonomics Science*, 3(2), 178-194.

462. Young, M. S. & Stanton, N. A. (2002c). It's all relative: defining mental workload in the light of Annett's paper. *Ergonomics*, 45(14), 1018-1020.

463. Zapf, D. (1995). An action theory framework of job characteristics. In J. M. Peiro, J. L. Melia, O. Luque, & F. Prieto (Eds.). *Work and Organizational Psychology: European Contributions of the Nineties* (pp. 61-76). Erlbaum, UK: Taylor And Francis.

464. Zapf, D., Vogt, C., Seifert, C., Mertini, H. & Isic, A. (1999). Emotion work as a source of stress: the concept and development of an instrument. *European Journal of Work and Organizational Psychology*, 8(3), 371-400

465. Zeier, H., Brauchli, P. & Joller-Jemelka, H. (1996). Effects of work demands on immunoglobulin A and cortisol in air traffic controllers. *Biological Psychology*, 42, 413-423.

466. Zimmermann, P. & Leclercq, M. (2002). Neuropsychological aspects of attentional functions and disturbances. In M. Leclercq & P. Zimmermann (Eds.), *Applied Neuropsychology of Attention* (pp. 56-86). New York: Psychology Press.

467. Zuckerman, M. (1979). *Sensation Seeking. Beyond the optimal level of arousal.* Hillsdale, NJ: Erlbaum.

INTRODUCTION AND OUTLINE

The field of aviation has undergone a lot of transformations since the first successful powered flight more than hundred years ago. Not only reflected in the technical progress, traveling with aircraft has become a major need for passengers from different backgrounds. The introduction of Air Traffic Services such as Air Traffic Control (ATC) was necessary to deal with the increasing traffic demands. Further initiatives towards unifying European airspace around the *Single European Sky* as laid down in the *Framework Regulation* EC 549/2004 of the European Community reflect the internationalization in this field. One of the major challenges for Air Traffic Management (ATM) concerns the organization of future traffic. STATFOR EUROCONTROL (2004*)* estimated that air traffic in Europe is going to double within the next 20 years. In order to cope with this expected growth, many concepts are currently developed to guarantee safe and efficient air traffic within these future scenarios.

For EUROCONTROL, "…solutions may exist only at the European level, by introducing new concepts into air traffic management and by favoring HOMOGENEITY" (EUROCONTROL Gate to Gate, 2005). Thus, it is not surprising, that the aspect of repetition is emphasized in the operations of the Central Flow Management Unit through anchoring the input of repetitive flight plans (EUROCONTROL CFMU, 2004). Conversely, a tendency in the opposite direction is observed in the International Civil Aviation Organization (ICAO), who defined flexibility as one of the key performance areas that promotes a more flexible use of the airspace.

Within these developments the consideration of an issue like monotony in a research program is not obvious. Historically, monotony represented an important concept in the field of work and organizational psychology since the beginning of the 20th century. Initially, the focus was set on problems related to mass production. One of the consequences of the introduction of automation were uniform and repetitive tasks which dominated the work activities and were perceived as monotonous. With an increased amount of monitoring tasks, which were dominated by the detection of rare events, monotony was also related to that kind of task. After an initial interest to describe the subject in ATC 30 years ago, the awareness of its relevance decreased again. Although some researchers continued to work on this issue, it has not received the same attention as research on stress or overload. One reason may be found in the unattractive position of researching such a topic in a world that appears to be dominated by complex and fast changes resulting in stress and requiring stress management. But especially the topic of automation evokes interplay between underload and overload and consequently the discussion of monotony and boredom. Moreover, progressing standardization takes the opportunity for an air traffic controller (ATCO) to create variety in his or her working environment. Finally, a state of monotony in human operators has multiple impacts within a short-term and long-term perspective.

With regard to these developments, the goal of this thesis is to explore monotony in the domain of air traffic control. The knowledge of factors which evoke and contribute to monotony is necessary for improving organizational, individual and task conditions and provides the basis for task execution in optimized work conditions.

To investigate these issues, three experimental studies have been conducted in laboratory and field settings which are described in the subsequent chapters of the present document organized as follows:

- Chapter 1 discusses the relevance of monotony for ATC.
- Chapter 2 describes the background of this research. It includes the description of key characteristics, defines its limits and presents models and concepts that might help to explain the development of monotony.
- Chapter 3 presents the research framework serving as a background for the study approaches and summarizes the research questions. Through the application of a simulation approach, the contribution of selected factors to monotony has been investigated in a preliminary and a confirmative study, which will be reported in Chapter 4 and 5 respectively.
- To validate simulation results, a field study is presented in Chapter 6.

An overall discussion will conclude this work (Chapter 7), which will not only challenge the current theories of monotony, but will also propose recommendations how to address this issue in ATC.

1. MONOTONY IN ATC: THE DEFINITION OF A RESEARCH PROBLEM

Although the introduced developments give rise to the idea that along with the ongoing traffic increases the main tasks for an air traffic controller have changed, this has not been the case yet. Still, the enroute controller has to ensure that aircraft can safely navigate in defined airspace, a so-called sector, according to flight plan. A radar monitor and communication equipment are the main tools supporting proper planning and acting in agreement with given regulations. She or he assumes the aircraft when it enters the controlled sector and keeps required separation between aircraft through eventually changing the flight path to avoid separation infringements. The controller provides information to aircraft and adjacent sectors and issues clearance orders. Depending on the phase of the flight, the flight is taken over from or handed off to a neighboring sector in the same center, a sector assigned to a different center, terminal control or approach control (APP). The former deals with aircraft in departure from or arrival to an airport. Approach controllers sequence aircraft into the most efficient order for landing or place them in holding patterns. Finally, in Tower Control the air traffic controller (ATCO) guides aircraft through landing and controls movements onto and off runways as well as around terminals.

Potential conflicts occur if the required separation between two aircraft is not maintained. According to regional requirements, additional functions might be introduced. Generally, a planning and an executive control position share the tasks within the sector. The planning controller (PC) observes the evolution of the traffic and coordinates with adjacent sectors in order to avoid separation conflicts and reduce the load of the executive controller. The controller at the executive position (EC) is involved in direct communication with the pilot and has the final responsibility for any undertaken action. Work organization and terminology differ slightly between countries, even though the International Civil Aviation Organization (ICAO) provides the background for standardization and regulations within air transportation. Despite the fast changes in the ATM domain, there is still a difference between more advanced states and regions that apply procedural or non-radar control. Also, the introduction of automation varies between Air Navigation Service Providers (ANSPs), as far as it concerns for example the transmission of data linked information or the replacement of paper flight strips with their electronic counterpart. These and other tendencies led to the characterization of the European ATM as a fragmented system regarding the division of decision-making operational units on a national level. A recent report has addressed the costs related to this approach (EUROCONTROL Performance Review Unit, 2006).

On an individual level, task analysis revealed that an ATCO monitors the traffic, checks the flight-related information, diagnoses and decides how to solve potential conflicts, and controls the implemented solution to reach the goals (Kallus, Barbarino, & Van Damme, 1997). This emphasizes the relevance of underlying psychological processes such as perception and information processing, attention, memory, problem solving, and decision making. A basic process in the controller’s work is to build the picture of the traffic situation in his or her mind, which has been described by Whitfield (1979) and Whitfield and Jackson (1982). But it shall be noted that Bisseret (1970) and Sperandio (1974) had already started the description of the *memoire operationelle* in the years before. Through the intense study of memory in French ATCOs these authors laid down the basis not only for the description of the *picture*, but also the currently booming concept of Situation Awareness (e.g., Endsley, 1996; Banbury & Tremblay, 2004). The picture “provides the basic understanding of the traffic scenario as a whole on which planning, scheduling, predicting, solving problems and making decisions depend, and also provides the basis for checking that instructions are being obeyed, that decisions are correct and that plans reach fruition“(Hopkin, 1995, p. 312). An accurate picture is dependent on the underlying mental model and the strategies stored with it such as trajectory prediction and comparison of altitude, time, or distance (Nunes & Mogford, 2003). These strategies allow an ATCO to rely on during conflict detection when directing the traffic flow.

Finally, it appears that up until now rather peripheral features in the preconditions for task execution have changed through the introduction of new tools, while the basic elements of task execution described above are still the same for an ATCO.

Probably due to the hypothesized relationship between operational errors, increased traffic demands and long-term effects on well-being and health, the relevance of the stress concept has been widely accepted and rather well researched in ATC (e. g., Smith, 1980; Melton, 1982; Zeier, Brauchli & Joller-Jemelka, 1996; Vogt, Adolph, Azan, Udovic & Kastner, 2002; Ming et al., 2004). However, a linear relationship between stress and operational errors has not been found. Moreover, operational errors seem to be occurring mainly in four work situations, that is, obviously, during high work load, after return from break (both described in Della Rocco & Cruz, 1999), after a traffic peak (Hagemann, 2000), and paradoxically, under low or moderate traffic load (Stager, Hameluck, & Jubis, 1989; Weikert & Johansson, 1999). It needs to be noted that the reported results may partly be due to an unconsidered statistical artifact because low or moderate traffic load prevail in many centers. However, it is difficult to explain the contribution of monotony. Currently no evidence exists that directly relates the occurrence of monotony to an increased incident risk. Besides, already 30 years ago Thackray, Bailey, and Touchstone (1975) started the discussion of concepts that might be relevant in this context in ATC. Nevertheless, Hopkin (1995, p. 341f.) needed to point out several arguments, why research on boredom - considered as closely related to monotony, is still important. In addition, the following findings underline his recommendations. Through analyzing the operational error database of the FAA, Schroeder (1982) revealed that most errors occurred in periods of low or moderate workload. Stager, Hameluck, and Jubis (1989) analyzed 301 operating irregularities in order to identify factors most likely to cause air traffic control incidents. About 80 % of the operating irregularities were found to occur more frequently under conditions of moderate or low workload and normal complexity. Probable causes were seen in problems of attention, judgment and communication, although no information on the ATCO' s perception of the situation was available. Similarly, Redding (1992) reported that 72 % of errors in 46 incident reports occurred with less than ten aircraft under control. He found that failures to maintain situation awareness (SA) cause the most errors in moderate traffic load, which was expressed in communication and coordination failures and the misuse of radar data. More recently, Weikert and Johansson (1999) investigated 36 Swedish incident reports and also found 25 out of 36 incidents occurring in low/moderate traffic density.

To conclude, it becomes obvious that the importance of low or moderate traffic load for air traffic safety has been underestimated. Several aspects have not been investigated yet and the issue of monotony merits further attention.

Overall, the problem can be approached from two perspectives. First, there are periods during which controllers have very little traffic to manage, a situation typically occurring in night shifts and frequently perceived as monotonous. Second, sectors and traffic flows have characteristics that might be defined as repetitive independent of traffic load. Both situations can be related to monotony, even though based on different causes. Hopkin (1995) also noted that temporal work structuring, expertise, personality, physical state, motivation, and job satisfaction are connected to monotony.

At the same time, solutions to mitigate monotony may apply to the operational environment as it is today as well as to the development of future concepts. But none of the current practices do support the understanding of monotony in either context. For example, common incident reporting systems do not systematically gather information about individual controller factors that might help to define monotony-related incident precursors for suboptimal individual states. In addition, models have not been developed that systematically explain errors under low traffic load. System designers maintain the trend towards further automation in the ATC environment in spite of a general awareness of critical side-effects and progress in the domain of adaptive automation (Wilson & Russell, 2003b).

Concepts like the synchronization of traffic flows are developed to cope with increasing traffic demands without considering that they might contribute to more uniform and homogeneous tasks evoking monotony. In such cases, even under high traffic density, monotony may occur because of the short action cycles in the task that reinforce the subjective feeling of monotony.

Till date, only few researchers addressed the phenomena related to *monotony* in ATC despite its obvious relevance. As a result, we have a limited understanding of this subject and its role in the performance of ATCOs. Therefore, the present thesis focuses on the development, evolution and management of the state of monotony through addressing individual (ATCO) and organizational (ANSP) perspectives. This knowledge will contribute to an increased understanding of factors related to performance and consequently to the safe and efficient handling of air traffic.

2. REVIEW OF ATC-RELEVANT LITERATURE AND THEORETICAL BACKGROUND

2.1. MONOTONY AND RELATED CONCEPTS: DISTINGUISHING AMBIGUOUS TERMS AND THEORIES

Well-designed work conditions are not only the basis for good performance of an ATCO, but contribute to the acquisition of new skills and self-confidence. Nonetheless, fatigue or other suboptimal states emerge during task execution. Monotony, as one such state, is introduced in this chapter. In a frequently used definition, monotony is seen as a “slowly developing state of reduced activation which may occur during long, uniform, repetitive tasks or activities and which is mainly associated with drowsiness, tiredness, decreasing and fluctuating performance” (ISO 10075-1, 1991, p. 2). As a high variety in definitions of monotony can be found, in this chapter the major focus is set on the distinction of related phenomena. From a state perspective, basically fatigue and satiation need to be distinguished, which can however just be achieved with the consideration of multiple assessment levels. The former addresses energetical aspects while the latter focuses on motivational aspects.

It is a common problem within psychology that concepts are described with a variety of terms or interpreted within different frameworks, as it occurs within research on stress. Also, within the context of monotony the application of ambiguous terms does not facilitate a consistent approach, since expressions like monotonous state, monotonous working conditions, underload, and boredom are used and often arbitrarily exchanged. For example, boredom has been used interchangeably with monotony, even though more recent publications discussed it in the context of emotional states (e.g., Kass, Vodanovich, & Callender, 2001). Different intercontinental research traditions have contributed to maintain this confusion. Despite the early interest in this topic in the 1920s, the number of scientific contributions remained small. In 1981, Smith presented an overview of the general research trends in the previous decades, where he indicated less than one publication a year. Only fairly recently an increased interest in this issue could be noted. Unfortunately Smith’s review focused on boredom research in the Anglo-American area and did not consider the literature offered by researchers from the European Continent (e. g., the German work group around Winfried Hacker, *cf.* e.g., Hacker & Richter, 1984) and the former Soviet Union (Gereb, 1968, 1978; Rzepa, 1984). Thus, the following chapter examines the phenomenology of related expressions and their common characteristics to clarify this confusion.

2.1.1. Monotony

When asking ATCOs about the meaning of the term *monotony*, they often answer that it is something that is repetitive, where it is always the same, something that can become boring. In the colloquial language, there is however no precise distinction of task characteristics leading to monotony and the individual consequences of exposure to these situations, the experienced monotony. While English texts often refer to monotony as a description of the situation, this definition might cause problems as it is not very clear. The Oxford English Dictionary (I, 1979, p. 628) defines monotony as the sameness of a tone or a tedious recurrence of the same objects. But there are also authors who recognized a difference between the evoking conditions and the consequences. In several publications Bartenwerfer (1957, 1961, & 1985) pointed out the importance of differentiating monotonous working conditions and an individual state of monotony. McBain (1970) recognized a discrepancy between monotony as a notion for the description of the environment, while at the same time it refers to individual consequences. According to the author, a stimulus situation is *monotonous* as long as it remains unchanged or changes only in a repetitive and predictable way. On the other hand *monotony* refers to the stimulus situation experienced by the individual, in its objective and measurable dimensions.

Similarly, Melamed, Ben-Avi, Luz and Green (1995a) distinguished objective and subjective monotony whereas Hopkin (1995) compared monotony with a subjective state of boredom and concluded that monotony is a task attribute as well as a subjective state that does not necessarily guarantee boredom.

Not as much attention has been devoted to the description of essential characteristics of monotonous work. Cox (1985) believed that repetitive work is only one form of monotonous work "…in which a discrete set of task activities is repeated over and over again in the same order without planned interruptions by other activities or tasks" (Cox, 1985, p. 86). In this approach, the cycle time for the set of activities is used as an index of repetitiveness. Alternatively, Bailey, Thackray, Pearl, and Parish (1976) discussed monotony as a condition of insufficient stimulation resulting from a lack of variety, interest or excitement that leads to boredom. The model of demands and load (Oesterreich, 1999) considered monotonous conditions as a certain type of load, in which it is required to continuously focus attention on certain information, events or conditions without the necessity to think, plan or decide. Characteristic examples are routine tasks or situations that contain long periods without the need to intervene. Another key term characterizing monotonous work is *uniformity*. Ulich (2001) distinguishes between temporal and content uniformity. The former has the potential of unburdening at work while the latter may foster monotony.

In contrast to rather objective descriptions of monotonous work, the state of monotony has been repeatedly described as a specific consequence of work strain (e. g., Richter & Hacker, 1998; ISO 10075). Bartenwerfer was the first to establish a theory of monotony (1957, 1961). He defined monotony as a state of unspecific physiological deactivation with reduced action readiness or capability. It develops in working situations that require continuous allocation towards restricted tasks, mainly characterized by low stimulation, high repetition, a low difficulty level, long time-on-task, and no possibility to change the task. A state of monotony is described on multiple levels. Subjectively, the task is perceived as uniform, boring and blunting; time is perceived as long; individuals show an unconcerned, apathetic attitude and resistance against the task; attention decreases, and a feeling of drowsiness emerges. The reduced ability to react and readapt to changing conditions results in impaired, varying performance. On a physiological level, a general deactivation is found, amongst others expressed in cortical and cardiovascular indicators. Hulin and Blood (1968) concentrated on the cognitive component when they defined monotony as the perception of the sameness of the job from minute to minute, with someone focusing on unchanging characteristics. Smith (1955) added that the perception of uniform or repetitive stimulus conditions induces a desire for change or variety. Unfortunately, the long tradition in monotony research has ignored that completely different task characteristics might evoke monotony. This was finally systematically considered by Johansson (1989) who distinguished *uneventful* and *repetitive* monotony. She compared control-room operators as an example of uneventful monotony with assembly line workers representing repetitive monotony. However, in her paper she adopts the term monotony to describe task conditions and thus maintains the unclear classification of monotony as cause and consequence.

Various attempts have been undertaken to explain the effects of monotonous work. In an early work McDowell and Wells (1927, quoted in Barmack, 1939a,b) postulated that the feeling of monotony is due to inadequate blood circulation where a fall or inadequate rise of blood pressure would accompany the feeling of monotony arising out of a bored attitude. Not confirming this assumption in his experiments, Barmack (1939a, b) suggested that the feeling of monotony is a result of the operation of more general factors such as the "tendency of the subject to revert to sleep, or a sleep-like state during the operation of a task-set" (Barmack, 1939a, p. 494). Apparently, the perception of fatigue in tasks that were originally used to evaluate boredom led Barmack to introduce the term "fatigue-like state". This becomes obvious in the statement that "a feeling of fatigue is (…) usually associated with boring work. Because of its transient character (…) it is generally referred to as a pseudo-fatigue or feeling of monotony" (Barmack, 1939c, p. 470).

An important concept to explain physiological deactivation in a state of monotony is habituation. O'Hanlon (1981) thought that monotonous sensory stimulation depressed the perceptual and cognitive functions of the cerebral cortex. He assumes that repetitiveness directly affects habituation while it inhibits cortical arousal. A compensatory process, effort, is elicited to restore arousal to an optimal level of task performance. When effort is no longer able to counteract habituation, cortical arousal declines below the point necessary for supporting acceptable performance. Finally, he concluded that the habituation hypotheses is not sufficient to explain low behavioral efficiency and argued that habituation occurs much more rapidly and usually follows a more monotonous time course compared to the performance decrement. Support for this argument is offered by Gillberg and Akerstedt (1998, *cf.* 2.4.1) who found that the effect of repetitive tasks on performance was already present at the beginning of the task. Desmond and Hoyes (1996) join O'Hanlon's discussion, when they assume a failure of effectively mobilizing effort under low task demands to explain a reduced average number of landed aircraft in low task demand compared to moderate or high task demand. A deeper analysis is undertaken in the context of action regulation theory, where monotony is considered as a consequence of a lack of sub-goals and fuzzy decisions concerning the goal-setting level as well as lacking variety in the task execution component (Rau & Richter, 1996). Oboznov, Yegorov and Kostritsa (1991) empirically investigated the role of goals in eleven operators and found that they transformed the mental image of task goals during the execution of a monitoring task. Operators who gave a personalized meaning to their success criterion (e.g., to prove good performance to someone), adequately retained its relevance for a longer period. However, during the task execution nonspecific activation procedures (e.g., pondering work problems, intensified motor activity) became a significant success criterion for operators. Even though these criteria were intended to support to achievement of the task goal, they turned out to be inadequate regarding the performance because of too much focusing on ones own condition. Moreover, the results indicate that self-regulation is not an ideal strategy for operators to maintain an optimal state, but individuals rather require external support.

Despite these relations with performance impairments and negatively perceived effects some authors also considered positive aspects of boredom and monotony (Bartenwerfer, 1985; Harris, 2000, Vodanovich, 2003). Monotony can enable relaxation and recovery, creative activities, search for change and variety and does not exploit all energetical resources during work, thus leaving energy for leisure activities. This was considered by Rzepa (1984). After investigating 357 workers she distinguished three types of post-monotony states. In post-monotony state I, operator-task-incompatibility is characterized by low reactivity and a preference for complexity accompanied by unpleasant, negative emotions, irritation and boredom with a decrease in mental functions. If operator capacity and task requirements are compatible, post-monotony state II emerges, where individuals show high reactivity and prefer simplicity. These subjects react with a calm working attitude, do not show negative feelings and evaluate work as easy and automatic. Post-monotony-state III is similar to state I, but individuals do make an effort to adapt to the requirements.

In addition, the discussion of routine, namely the unvarying or habitual procedures in task execution, is taken into account in ATC. Repetitive control strategies foster preset solutions that can be reapplied and thus contribute to an acceptable level of workload, since decision and reorientation processes do not need to be undertaken repeatedly. Therefore, routine tasks also relieve the burden, less cognitive resources are used and energy is saved. On the other hand, routine might encourage so-called mental sets or mechanization of thought. When people are biased by previous experience to prefer certain approaches to a problem, it may block the solution in a particular case (Luchins, 1942). One example is the maintenance of a mental picture or a strategy without recognizing the availability of a more effective solution. In this case, an update of the information presented does not occur and leads to an inappropriate mental representation of the situation. Transferred to the field of ATC it would mean that ATCOs keep the routine procedures without noticing that the situation has changed. There, also, the aspect of predictability plays a role.

Predictability is essential for planning, as it regards the immediate time horizon but also relates to experiences in the past. Only when characteristics of a situation are recognized again, it might be perceived as repetitive and, consequently, as boring. Even though not specifically addressed in ATC, this aspect might contribute to incidents, as repetitive, homogeneous and well-known traffic flows lead to complacency and overconfidence or to the expectation that the situation does not change.

Overall, the different aspects that came up in the description of monotony need to be carefully considered when researching this issue. For the current work, *repetitiveness* and *uneventfulness* are the preferred labels to designate task conditions. Compared to *monotonous* the use of these terms has the advantage of more clearly characterizing task characteristics and avoids implicitly expressing subjective valuations. On the other hand, *monotony* will be applied to the consequences evoked by such task conditions on an individual level such as reduced psychophysiological activation, sleepiness or impaired performance and clearly designate a state. Its relevance for ATC is indubitable.

2.1.2. Boredom and Underload

Various authors recognized the arbitrary and interchangeable use of the terms of monotony and boredom in the Anglo-American research area (Thackray, 1981a,b; Davies, Shackleton, & Parasuraman, 1983). Consequently, a higher number of research reports in the English language has been published on boredom than on monotony, while the German research tradition focused on the concept of monotony. Some arguments will be pointed out in the course of these chapters that are in favor of a clear distinction between monotony and boredom at work. As Fisher (1993) argued, everybody experiences boredom from time to time. The reasons to consider boredom as a risk-factor in ATC can be explained by the strategies people use to counteract boredom. Daydreaming, motor restlessness, exploration, and withdrawal (Harris, 2000, p. 581) are behaviors that might distract ATCOs from their primary task and thus have an impact on performance. In general, it needs to be distinguished if the task characteristics (something is boring) or the personal state (someone is bored) are described. Baldamus (1951) suggested that the effects of monotony and repetitiveness should be distinguished from content-boredom.

In most approaches, and as early as Hoche in 1923, boredom has been conceptualized as an affective construct. It has been defined as “a negative, dissatisfying emotional state” (Kass, et al., 2001, p. 319) or “an unpleasant, transient affective state in which the individual feels a pervasive lack of interest in and difficulty concentrating on the current activity” (Fisher, 1993, p. 396). Mikulas and Vodanovich (1993) discussed the essential aspects of boredom, which they defined as a state of relatively low arousal and of dissatisfaction, attributed to an inadequately stimulating situation. A general negative affect or unpleasantness was addressed in an experiment of Geiwitz (1966). He presumed that repetitiveness, constraint, arousal, and unpleasantness are essential factors of boredom. To investigate their relationship, four students executed a simple repetitive task (making checks on papers) in permutated, repeated measurement conditions. In each condition, one of the four factors was manipulated while the other three were kept constant. The results revealed that reported boredom is associated with low arousal, increased feelings of unpleasantness, constraint, and repetitiveness. Repetitive tasks may have an effect on boredom by decreasing cognitive arousal rather than by increasing subjective repetitiveness. He assumed that this effect emerged from the task’s potential to induce all four factors, whereas none of these factors were the *sine qua non* of boredom. However, due to methodological weaknesses, e. g. the induction of boredom through posthypnotic cues or inappropriate statistical analyses procedures (high alpha-inflation due to many correlations) no clear conclusions can be drawn. Later on, Scerbo considered the aspect of constraint (2001) when stating that boredom arises as soon as individuals are required to work at a task with highly repetitive, homogeneous stimuli beyond the point where they would normally reject it.

The connection to arousal was established in the model of emotional space along the two dimensions of valence and arousal, as it was already described by Wundt (1886) and discussed more recently in Larson and Diener (1992). In this model basic human emotions are located on a circle with tired/bored situated between relaxation and unpleasantness. Experimentally, Bailey et al. (1976) examined the relationship between boredom and arousal in 32 male participants performing a task of either high or low visual complexity. They found a complex response pattern in physiological and subjective measures that did not clearly indicate increasing or decreasing arousal. A possible explanation is that the feeling of boredom is a concomitant addressing the emotional component of a state and can thus be hardly induced in an experimental design. In consequence, the aspect of low arousal in the definition offered by Mikulas and his colleague cannot be supported. O'Hanlon (1981) presents a concept of boredom to explain why stimulation repetitiveness is a determinant of boredom and defined boredom as "a unique psycho-physiological state that is somehow produced by prolonged exposure to monotonous stimulation" (p. 52). His review also considered long-term effects seeing that social and medical consequences can be related to a repeated elicitation of the acute psychophysiological state which may constitute a condition of disturbed homeostasis

Another explanation for boredom was offered by Czikszetimihalyi (1975, 1993) in the context of flow theory. He argued that mood states are affected by the balance of activities or tasks between challenge and skill. Individuals are likely to experience boredom when skills exceed the challenge of an experience. When high levels of challenge are met with high levels of skills, a state of flow occurs. In this state of optimal experience, attention is focused on the activity, one is absorbed, loses self-consciousness, feels in control of the environment and the focus of awareness is narrowed. Originally four flow components were identified (Czikszetimihalyi, 1975) and extended to eight dimensions (1993): clear goals and immediate feedback, equilibrium between the level of challenge and skill, merging of action and awareness, focused concentration, sense of potential control, loss of self-consciousness, altered sense of time, and experience becoming autotelic or self-rewarding. In summary, according to the publication in 1993 the quality of experience can be estimated under consideration of a hedonic aspect, arousal, cognitive efficiency and motivation. Unfortunately, there are several problems with this concept. First, the proof of his theory is currently still insufficient and the distinction of the dimensions is problematic. Second, methodological weaknesses apply for the developed instruments. The Experience Sampling Method (EMS) was designed to overcome disadvantages of retrospective methods but interrupts respondents at random intervals, which might have a disruptive effect on an otherwise positive mood state. Also, no clear definition for boredom was offered. Overall, the consideration of a state of flow might be highly relevant in the discussion of monotony mitigation. Csikszentmihalyi and Le Fevre (1989) found in a field study of 78 adult workers that flow experience was reported more often during work than during leisure. If highly motivated, this experience was even more pronounced. It is however interesting that Czikszetimihalyi was not the only one to work on that issue. A very similar concept is found in the description of traction as undertaken by Baldamus (1961, quoted in Davies et al., 1983), which is a "feeling of being pulled along by the inertia inherent in a particular activity", accompanied by pleasant experience. This can be bound on an object, a batch of articles, a process, a machine or the line of objects passing along while working on them and thus counteract boredom.

The gap between boredom and underload was filled when Welford (1965) defined boredom as a state where the organism is underloaded. In the proposition of an underload/overload continuum McGrath (1976) integrated underload as opposed to overload to describe inadequate job demands and underutilization of skills. Here, a direct link to the description of boredom in Czikszetimihalyi can be built. However, quantitative and qualitative underload in the task is just one possible cause of job boredom, as Fisher (1993) summarized after a survey in 1987. Other major causes were qualitative overload because of excessively difficult tasks, the absence of colleagues and organizational constraints. Therefore, it is preferred to focus the definition of underload on individual abilities and needs, which continuously fall below those required by the task.

Such a definition does not equate underload and boredom and disclaims any valuations referring to the individual consequences. Already in 1960 Ulich suggested underload as a problem that work psychology has to deal with. Richter and Hacker (1998) distinguish quantitative underload where task demands are too rare from qualitative underload that refers to frequent, but uniform tasks without sufficient engagement. In terms of Johansson (1989), uneventful monotony may be seen as equivalent to quantitative underload and repetitive monotony to quantitative underload. In the opinion of Ulich (2001) boredom results from quantitative or qualitative underload, while monotony results from the feeling that one needs to do always the same thing; both can be related to feelings of fatigue.

An important contribution came from Hill and Perkins (1985) who focused on the cognitive component. They defined boredom as a subjectively perceived state and assigned a cognitive, an affective and a physiological component. The cognitive component describes the perception of task characteristics while the affective component describes how these characteristics are interpreted. The authors assumed that people construct tasks in a variety of manners and associate a wide range of constructs and distinctions and consequently the task is perceived as manifold and people are interested in the task. On the other hand, when people perceive tasks as homogeneous and undifferentiated, they connect monotony and boredom and combine it with frustration. They confirmed their assumptions in four experiments (n=92) using a repertory grid technique, which is based on the assumption that individuals interpret the world according to their own set of constructs (Perkins & Hill, 1985). Subjects who spontaneously used more constructs to describe objects and made finer distinctions on rating them were less bored. Also, they found that physiological changes such as increasing heart rate variability can - but not necessarily do - accompany boredom (*cf.* Chapter 2.4.1). So, boredom occurs when stimulation is construed as subjectively monotonous and when few constructs are applied, with the result that a high level of frustration is experienced. In this light, it might be assumed that if deactivation is a consequence of habituation in monotonous tasks, as O'Hanlon (1981) discussed, interested people should also show deactivation. Unfortunately no conclusion can be drawn, as he did not report any remarks on people's mental construction of the situation. Besides, attention to one's internal states and self-reported affective involvement mediate the experience of boredom (Swinkels & Giuliano, 1995).

As pointed out by Smith (1981), boredom was not only of interest in working environments but relevant contributions were made in the psychiatric field. For example, Fenichel (1951) distinguished between existential boredom as a source of chronic suffering from boredom that is dependent on the situation and environment. Bernstein (1975) differentiated between chronic boredom as a chronic feeling state and responsive boredom as an affective response to a certain external situation. Based on psychoanalytic theories, Revers (1949) described boredom as a form of apathy, not finding any interesting object or task to reduce the tension in one's drive. Dynamic boredom refers to the immediate loss of interest in any object or task, whereas hectic boredom comes up in a situation where any personal commitment is avoided and someone opposes meaningless situations with absolute indifference. Similarly, Berlyne (1960) regarded boredom as a drive that is reduced through divertive exploration and aroused when external stimuli are excessively monotonous.

Nonetheless, in all these cited definitions boredom was somehow seen as a temporary condition while the description of boredom as a trait remained sparse for a long time. However, as already announced in the term *existential boredom*, individuals were identified who had a propensity to be bored across time and situations. Therefore, Zuckerman (1979) included boredom susceptibility as a subscale in the Sensation Seeking Scale, where it is defined as an "aversion for repetitive experience of any kind, routine work, or dull and boring people, and extreme restlessness under conditions when escape from constancy is impossible" (p. 103). At this point the similarity of this definition to the later described concept of satiation is noted. Farmer and Sundberg (1986) developed the Boredom Proneness Scale which addresses "one's connectedness with one's environment on many situational dimensions, as well as the ability to access adaptive resources and realize competencies" (p. 10).

Their multifaceted profile of the boredom-prone individual includes distractibility, depression, dissatisfaction with work, low motivation, and a lack of autonomy.

As it becomes obvious, there is no general agreement on what boredom is and where differences to monotony and underload can be found. Often boredom and monotony were thought to be a consequence of the same situation, but this does not always hold true. For example, monotony as a consequence of task characteristics may result in demanding situations, at the same time individuals experience fatigue. Summarizing all endeavors to define boredom as a consequence of understimulation in the words of McBain (1970): "boredom may or may not result from monotonous work conditions" (p. 509f.). The concept of boredom is not just present in situations where the environment is judged as repetitive or uniform, but it seems to be important in any situation where an affective reaction occurs, as people are different in the manner in which they cognitively construct and interpret a situation. However, the classification of boredom as an emotion in the model of Larson and Diener (1992) is difficult when current discussions in the field of emotion psychology are considered (e.g., Plutchik, 1980). There, it is clearly distinguished between emotion and mood states. Emotions are described as short-lasting intense reactions to events deemed relevant to the needs, goals or concerns of an individual. In contrary, mood is experienced as more diffuse, global and general. It may be indirectly caused by a particular object but is not directed by any object and seen as a long-lasting general affective state (Linnenbrink & Pintrich, 2004, p. 58). In consideration of these developments the described phenomena in underloading task conditions would rather denominate boredom as a mood state than an emotion (*cf.* 2.3.3 for a distinction between both).

2.1.3. Fatigue

Fatigue is a term used to describe many different experiences like sleepiness, tiredness, or physical exhaustion. As a complex state, fatigue overlaps with areas of performance, cognition, physiology, emotion, and also with boredom and drowsiness (McDonald, 1989). But, as Michielsen, Vries, Van Heck, Van de Vijver, and Sijtsma (2004) stated, "not much systematic theorizing has taken place yet" (p. 39) and definitions of the construct are poorly described in most of the current fatigue studies.

First of all, it is indicated to differentiate various forms of fatigue. Richter and Hacker (1998) distinguish fatigue as a consequence of the circadian rhythm from fatigue as a consequence of task execution that results in a reversible reduction in performance of an organ (local fatigue) or the whole organism (central fatigue), or can be peripheral (physical) or central (mental) fatigue (Gawron, French, & Funke, 2001). Time-on-task was proposed as a better term to use for task-related fatigue (Van der Hulst, 2001), since fatigue also can be chronic and develops over time (*cf.* section 2.2.2). In the definition of Soames-Job and Dalziel (2001) fatigue refers to the "state of an organism's muscles, viscera, or central nervous system, in which prior physical activity and/or mental processing, in the absence of sufficient rest, results in insufficient cellular capacity or system wide energy to maintain the original level of activity and /or processing by using normal resources" (p. 469). Even though this definition includes a more precise description of physiological processes, the authors do not include any statement concerning its assessment. In the research group around Meijmann (e.g., Van Dijk & Swaen, 2003) fatigue is seen as the change in the psychophysiological control mechanism that regulates task behavior, resulting from preceding physical and mental efforts, which have become burdensome to such an extent that the individual is no longer able to adequately meet the demands that the job requires on his or her mental functioning or only at the cost of increasing mental effort or mental resistance. Therefore, reduced competence and willingness to develop or maintain goal-directed behaviors aimed at adequate performance is found. Hence, the feeling of fatigue might also be considered as a stop-emotion (Gaillard, 2003).

Fatigue is not only task-related, but emerges in the context of circadian biorhythmic sleep-wake-cycles, which have a big impact on human performance (Costa, 1999). Physiological and psychological systems follow a certain temporal organization, wherefrom circadian rhythm is mainly reflected in body temperature. Since most of the individuals dispose of a regular sleep/wake-cycle, sleep deprivation or a loss thereof is one of the factors resulting in impaired performance that has been considered in many studies and is especially a problem in shift work (*cf*. section 2.4.2). Still, shift work is seen as a major issue in the context of fatigue, as sleep disturbances are reported more often when one tries to work against the circadian rhythm (Rosekind, Gander, Gregory et al., 1996a,b). For this reason, it has been a subject of interest (Rosekind, Gander, Miller, et al. 1994) especially in the aviation domain. In this context the finding of Spencer (1997, quoted in Gander, 2001) is interesting. In a diary study, fatigue-ratings remained stable under low workload conditions for up to four hours, while a rapid increase was observed after two hours when workload was high.

The multidimensionality of the perceived fatigue was researched by various authors, but Ahsberg (2000) found *lack of energy* as a general latent factor in the prior dimensions of physical exertion, physical discomfort, lack of motivation and sleepiness. Bartenwerfer (1961, p. 253) emphasized that the term fatigue has to be used when talking about changes in the psychophysic structure. He stressed that the process of strain leads to a state of fatigue and assumed the involvement of the central nervous system (CNS) on a sub-cortical level, which might explain reduced vegetative functions.

An alternative approach to define fatigue concentrates on the assessment of indicators like the quantity and quality of performance, self-reports of fatigue, sleepiness, weariness and dislike of work, disruption in reception and perception, coordination, attention, concentration and social relations (see Grandjean, 1991, p. 163f., Richter & Hacker, 1998, following Schmidtke, 1965; see also Luczak, 1998, p. 280). On the physiological dimension decreased heart rate (HR) and respiration amplitudes, increased heart rate variability (HRV), alpha-waves in electrocardiogram (EEG), blink rate and flicker frequency were reported (further described in studies below). Subjectively, increased tiredness and reduced concentration were reported (Weber, Jermini & Grandjean, 1975). But as it became already clear, subjective reports are not sufficient to characterize fatigue since these feelings can also be related to boredom or monotony. In an air traffic control center the development of fatigue was investigated during three weeks. It was found that flicker frequency and the scores in performance tests decreased, and fatigue and sleepiness were reported more frequently after seven hours of work (*cf.* Grandjean, 1991 p. 170ff.). Gregory, Oyung and Rosekind (1999) analyzed 153 fatigue-related ATC incident reports (2,7% of total reports in eight years in a voluntary reporting system) and defined controller fatigue as the most often identified factor followed by workload and duty factors, but incidents were not related to night shift or lighting conditions. Morris and Miller (1996) investigated the relationship between flight simulator performance and oculometric measures in ten pilots after a night of sleep deprivation and found blink amplitude, blink rate and long closure duration as the best predictors for performance degradation due to fatigue. However, as Caldwell and Ramspott (1998) showed, task duration is a significant factor to demonstrate the effect of sleep deprivation. Dinges, Pack, Williams, et al. (1998) also showed the effect of continuous sleep restriction on performance and subjective well-being that continued beyond the completion of the task. No effect of sleep deprivation was found in a primary task performance, but it resulted in strategy changes and subsidiary task impairment in a machine-centered process control task executed by 16 participants (Hockey, Wastell, & Sauer, 1996). The onset of fatigue was also observed in a secondary task while primary performance remained unaffected (Mascord & Heath, 1992). Strategy-based rather than capacity-based changes in performance were the favored interpretation by Monk and Leng (1982) in the investigation of time-of-day effects.

Van der Linden, Frese, and Meijman (2003a) tried to explain the mechanism of how fatigue impairs performance. It was assumed that mental fatigue affects control processes involved in the organization of actions and plays a major role in deliberating goal-directed behavior. If goal activation is reduced, actions are guided by more automatic processes. They confirmed their assumptions with 58 students exposed to fatiguing or non-fatiguing tasks for two hours. After this introduction they had to judge their motivation, mental effort, mood and intelligence. Afterwards, the Wisconsin Card Sorting Test and the Tower of London were used to assess the impact of fatigue on executive control. In the fatiguing condition, preparation processes as indicated in event-related potentials were impaired and increased errors were found. In another study using a similar design with 68 students the influence of low and high experience was also considered (Van der Linden, Frese, & Sonnentag, 2003b). Fatigued participants with less computer experience showed more rigid behavior, which was explained by reduced task engagement. In addition, Lorist, Klein, Nieuwenhuis, De Jong, Mulder, and Meijman (2000) found that brain activity is reduced with increasing time-on-task in areas of the frontal lobe that are associated with the exertion of executive control. Alternatively, so-called *blocks,* which are interruptions of information processing, were the interpretation offered by Schroeder, Touchstone, Stern, Stoliarov, and Thackray (1994) when they found impaired conflict detection in a two-hour-simulated ATC task executed by 20 students during three days. As a symptom of fatigue, it had already been discussed by Bills (1931) who described that it is not possible to concentrate on mentally loading tasks continuously and thus blocks occur more frequently.

Based on these studies the approach is challenged to define monotony as an independent construct. Desmond and Hancock (2001) considered separate active and passive fatigue states and integrated them in a framework of adaptive attention to explain fatigue. Fatigue occurs in a state of reduced attentional capacity to maintain the normal oscillation between sampling the environment and self-evaluation. This reduction occurs because of continuous activity, resulting in active fatigue, or as a consequence in chronic understimulation in passive fatigue. Already Kraeplin (1903, quoted in Bartenwerfer, 1961, p. 251[1]) pointed out not to confuse a feeling of tiredness with fatigue. For him a feeling of tiredness arises with any increased effort in a task. Moreover, motivation is an additional factor that contributes to explain why performance impairments are more likely to occur under high fatigue. And as mentioned by Hopkin (1995), trying to resist monotony is also fatiguing.

Prior studies do not demonstrate a clear indication of the distinction between fatigue and monotony. As Bartenwerfer described, there is always an alternating change between fatigue and monotony. Also, it might be possible that fatigue and monotony are just distinguished concepts at an early stage in the work, while with increasing time-on-task the concepts are approaching each other or overlapping. Finkelman (1994) investigated the database of a temporary employment agency and collected work-related information as well as reported fatigue. Comparing ratings of 3705 employees who experienced fatigue with 10000 randomly selected employees without fatigue indication, he found low job challenge, poor-quality supervision, poor job performance and low pay rates associated with subjectively experienced fatigue. Interestingly, positions with low physical demand and low information processing were also associated with subjective fatigue. In this light, results of many studies can be explained in terms of monotony or boredom as well as fatigue. That fatigue and monotony are different states can be supported by the observation that well-rested participants also experienced monotony very soon (Bartenwerfer, 1957).This process occurred even faster if the initial level of fatigue was already very high. Barmack (1939, p. 470c) also found that subjects rated monotonous tasks predominantly as boring, not fatiguing. In the same line Gereb (1978) refers to monotony as a pseudo-exhaustive state. On the basis of Signal Detection Theory (SDT) Frishman (1990) developed an experimental paradigm to investigate the visual discrimination efficiency in a 60 min discrimination task of different complexities, in varied order, to distinguish states of monotony and fatigue.

[1] Original document not available.

He found that in both a state of monotony and fatigue the discrimination ability reduces as indicated through a reduced probability of hits. On the other hand, the decision criterion strictness, indicated through increased probability for false alarms increases in monotony, but decreases in a state of fatigue.

An argument that supports the distinction between fatigue and monotony is given by Richter and Hacker (1998), who stated that these states have different causes and therefore need to be mitigated or avoided in a different way. Monotony can be eliminated immediately with variations in the job. Dueker (1931) slowly increased the pace in work what led to less monotony and thus supported the significance of task changes. This would favor the statement that the mechanism behind the states of monotony and fatigue might be different. Increasing pace of work accelerates the performance decrement in fatigue, while it leads to increased briskness in monotony. It is difficult to assess if monotony and fatigue are different states since they show similar characteristics and might be potentially overlapping. As a consequence, physiological and subjective indicators are not sufficient to define monotony, but in addition performance should be positively affected by alternating tasks.

2.1.4. Satiation

Another inadequate consequence of task execution already discussed at the beginning of the last century and mostly neglected since then, is the concept of satiation introduced by Karsten (1928).

Satiation is described as a state of increased tension, if someone feels agitated, annoyed, affect-laden, and not being able to move from a certain place. It is a situation in which a person does not want to continue to work on a task but has the obligation to do so. That can be in situations with repetitive tasks (Berman, 1939a) or any task (Ryan, 1947, quoted in Gubser, 1968). Ulich (2001) added that the attitude towards the task is more important than the repetition and Richter and Hacker (1998) mentioned satiation when there is low incentive and a person is not able to meet the demands.

Berman's definition (1939b, p. 281) that a satiated person is a person that rejected an object or activity was criticized by Barmack (1939c) due to an obviously arbitrary exchange of satiation and boredom. In a reply to Berman, Barmack (1939c, p. 469) saw the difference between their experimental studies on the concepts of boredom and satiation in the possibility to freely stop a task against which someone developed a negative valence when one is satiated, whereas in his own experiments the subject is acting under the constraint to complete a specified activity. From his experiments, he sees the psychic satiation as one probable aspect of the state of boredom. Boredom he sees as:

> " …a state of conflict between the tendency to continue and the tendency to get away from a situation which has become unpleasant principally because of inadequate motivation resulting in inadequate physiological adjustments to it. A state of boredom is initiated by inadequate motivation during the operation of a task set and results in a tendency for the physiology of the subject to revert back to the sleep level. The inadequate vital adjustments to the task are unpleasantly appreciated as the feeling of monotony or fatigue. If the task set is weak, the subject may go off to sleep or abandon the task. If the task set is sufficiently strong, the subject struggles to remain awake or partly escape from the depressing task. These later objectives are achieved usually unconsciously, by shifting attention away from the task, daydreaming, creating extrinsic goals, modifying the procedure, etc." (Barmack, 1939, p. 468).

However, this definition does not seem to be sufficient, since there are strong similarities with the later described concept of *satiation,* apparently representing the opinion of Berman (1939b). But Berman also did not cite the complete description of satiation as it was proposed by Karsten. Unfortunately, after these intense discussions the satiation concept was neglected in American research until Scerbo and his colleagues reintroduced it when researching boredom. The *Task-related Boredom Scale* (TBS; Scerbo, Rettig, & Bubb-Lewis, 1994) addresses eight factors that are thought to contribute to feelings of boredom: stress, irritation, relaxation, sleepiness, alertness, concentration, passage of time, and satiation. In addition, respondents estimate their overall feeling of boredom. Nevertheless, the disadvantage of this scale is the confusion around different concepts, even though single items might be appropriate.

A different focus was set in recent German developments (Schultz-Hardt, Rott, Meineken, & Frey, 2001). Therein, satiation is defined as the loss of intrinsic motivation to continue a task that results from the continuous repetition of the task. In 66 high-school students it was confirmed that if a repetitive task was of high personal importance and can be hardly executed without investing a lot of attention and rather peripherally to the main task, mental satiation was more likely to occur (Schultz-Hard, Meineken, Rott, & Frey, 2001). But the motivational component is not really new, as it was already considered in the definition of boredom (Barmack, 1939a, p. 495). In the light of Berman's conclusion that "attempts to indicate that all aspects of satiation cannot be explained on the grounds of basic motivational states, and tendencies to revert to a sleep level" (Berman, 1939b, p. 472f.), the conclusion of Schultz-Hardt, Rott, and coworkers (2001) is remarkable. For these authors the continuous task execution leads to higher accordance between structures of a person and the environment and increased familiarity. In consequence, a loss in intrinsic motivation emerges that is called psychic satiation and which is perceived as a loss of interest in the task.

The distinction between satiation and monotony is challenging as the same working conditions can lead to monotony or to satiation (Bartenwerfer, 1985); also, boredom seems to be related to both. It appears that this state is not just a consequence of task characteristics but also of a certain inner attitude of a person. As a defining component in satiation, tension should go along with increasing arousal while in a state of monotony tension should be decreased as indicated in deactivation. This dissimilar development might explain some of the contradicting results around monotony. It is also remarked that *satiation* may not be the most appropriate term to translate this concept that Karsten introduced as *Saettigung*. *Saturation* is another term which was used by researchers from different contexts who accessed the original literature (e.g., Gereb, 1978). Originally translated as satiation, an etymologic analysis does not really help to clear up this issue (Oxford English Dictionary II, 1979, p. 118). S*atiation*[2] is hardly ever used and if it is, then in a positive context. On the other hand *saturation* means that someone is physically full, that it is not possible to take anything more, for example as applied to a nervous system or a brain or a desire or appetite. To avoid confusion around the concept, the application of the first applied translation *satiation* is maintained.

2.1.5. Low vigilance

One concept that is important to reflect on in the context of monotony is *low vigilance*. In the internationalizing procedure for the International Standards for Mental Workload (ISO 10075-1) low vigilance was added to monotony. At first sight the vigilance concept appears to be rather well described, even though different conceptual approaches need to be considered.

[2] An interesting anecdote refers to how the concept was spread. It is hypothesized that thanks to Kurt Lewin the satiation concept was acknowledged in American research. After editing and contributing to the publication of Karsten (1928) he spent several months as a visiting professor in the United States before he finally emigrated. During his influential work various opportunities might have allowed to meet Berman or Barmack, who continued to work on the ideas of satiation.

In the tradition of Mackworth's basic work on radar operators (1948) many authors define it as the ability to maintain a state of readiness for a long time in order to detect and respond to specified infrequently occurring events in a stream of irrelevant events (Koella, 1982; Sawin & Scerbo, 1995). While this definition centers the performance aspect, other authors related vigilance to the capability of sustaining a certain level of cortical alertness, which follows the physiological approach of Head (1923), or equate it with sustained attention (Parasuraman, 1984). At the same time Makeig and Inlow (1993) criticize that many studies define vigilance simply with physiological criteria such as EEG and electrooculogram (EOG) without considering performance fluctuations. Thiffault and Bergeron (2003a) pointed out two broad conceptions of vigilance. One focuses on the physiological mechanisms related to activation and alertness as demonstrated in wakefulness and arousal, the other relates cognitive processes pertaining to one's ability to maintain sustained attention in a task. The ability to remain vigilant in terms of sustained attention fluctuates with physiological alertness that varies according to endogenous and exogenous or task-induced factors. As a conclusion the authors find that vigilance combines both alertness and attention, but do not further analyze the role of task goal-related aspects.

There are however difficulties regarding the various aspects of vigilance definitions. To equalize vigilance with sustained attention is not appropriate, since recent studies have shown that these concepts which are both supposed to represent the intensity aspect of attention (Van Zomeren & Brouwer, 1994) do affect different regions of the brain **(**Zimmerman & Leclerc, 2002**). W**hile tasks requiring vigilance are characterized by low information and rare targets, sustained attention requires continuous processing of a higher amount of information (Leclercq, 2002). The relation to alertness is also manifold. Posner and Rafal (1987) distinguished tonic and phasic alertness with the latter directing the attentional focus to an unexpected stimulus or event while tonic alertness incorporates fluctuations mainly related to circadian rhythms and thus wakefulness.

To investigate vigilance, variations of the original clock test of Mackworth (1948) were applied. The main results of decades of vigilance research are available in several reviews (e.g., Wickens & Hollands, 2000). One of the main findings is the vigilance decrement that designates a drop of accuracy or increased reaction time in the detection of a target signal. The vigilance decrement was found to be a function of time-on-task, signal frequency and intensity, knowledge of results, and many other endogenous and exogenous factors such as age (Deaton & Parasuraman, 1993). Different explanations have been proposed for the performance decline. Robertson, Manly, Andrade, Baddeley, and Yiend (1997) thought that the repetitive nature of vigilance tasks leads to mindlessness, as automaticity, routinization and lapses of attentional focus may result in the withdrawal of effortful attention away from the task. Support was found in a study where participants with a high score in a cognitive failure scale performed more poorly. It is noted that this explanation recalls Luchins' set-effect (*cf.* 2.1.1) and reflects an endogenous modulation of attention rather than the decline in wakefulness and vigor accompanying lowered arousal (Dickman, 2002). In contrast, Grier et al. (2003) confirmed in a modified vigilance task that vigilance decline might be better characterized by effortful attention (mindfulness) than by mindlessness because of cognitions that were involved. This is also close to the explanation that mental effort affects vigilance because of resource depletion (Smit, 2004a).

To characterize suboptimal vigilance, expressions from literature like *loss of vigilance*, *reduced vigilance, low vigilance* or *hypovigilance* have been applied. Hagenmeyer (2005) defines hypovigilance as a state of diminished vigilance that is often referred to as fatigue or drowsiness. Similarly, Muzet and Roge (2003) see low vigilance as a state that is reflected in the physiological state of drowsiness and progressively increases with time, but differentiate it from a sudden loss of attention. They propose employing this concept in situations where operators have to face a long and monotonous task. Nonetheless, the described phenomena are similar to those in a state of monotony.

Kirwan (2005) used the concept of low vigilance in ATC and referred to it as a decrease in controller's awareness that is related to fatigue, time of day, and low workload. However, this definition is not very systematic either as it does not distinguish the description of the task from individual consequences. Moreover, it contains concepts of a very different nature. Ulich (2001) saw reduced vigilance as well as monotony as results of a situation with a low number of stimuli necessitating bound attention, where vigilance is the result of a situation with a very rare requirement for reaction.

The inclusion of reduced vigilance in the International Standard for Mental Workload (ISO 10075) maintained the confusion. Therein, it is defined as a state “with reduced detection performance in monitoring tasks offering little variation” (ISO 10075-1, p. 2). However, the next paragraph continues with “(…) monotony and reduced vigilance can be differentiated with respect to the circumstances of their causal conditions, not with respect to their effects“ (ISO 10075-1, p. 2). Such a definition contradicts the initial purpose of the standards to facilitate the identification of different critical states as a consequence of work strain. It is preferable not to apply monotony and low vigilance as task descriptions next to the other critical states, but to clearly distinguish between the task and the individual reaction. Thus, it is favorable to deploy the terms uneventful and repetitive as a description of task features, while monotony is reserved for the state denomination. This is in agreement with the arguments of Richter and Hacker (1998, p. 118), who stated that performance in vigilance has to be seen in the light of states of monotony. Monitoring tasks fulfill the definition requirements set by Bartenwerfer, that the uniform task does not allow any distraction from the task nor to deal with the task. Another common component is the *readiness to action*, mentioned as a characteristic of monotony by Bartenwerfer and similar to the *state of readiness* in vigilance.

In ATC the vigilance concept is able to predict difficulties in unforeseen actions during an uneventful work situation. Nonetheless, there are several restrictions in applying this concept to this field. Mackie (1987) criticizes that in vigilance research a lot of attention has been put on factors having very little relevance for the operational field according to the judgments of 212 sonar operators. On the other hand factors such as boredom, monotony, fatigue and tiredness were judged highest. Furthermore, it is an artificial construct compared to monotony and boredom. Johansson, Cavalini, and Petterson (1996) state that the generalization from vigilance experiments to process monitoring is limited, as vigilance tasks do not contain dynamic sequences of events. Also, performance measures are attained from the vigilance task itself but not from unpredictable cognitive tasks. Still, a lot of experiments have been undertaken with ATC-related tasks (e.g., Schroeder et al., 1994). While this is true, there are numerous reasons noteworthy for not allowing a direct comparison of ATC with a classical vigilance task based on the required activities:

- Task analysis revealed that ATC consists of a variety of subtasks that are completed in addition to monitoring (e.g., Kallus, Barbarino, Van Damme, & Dittman, 1999). These include planning the expected traffic or updating one’s picture about the traffic situation through scanning the screen even if there is little traffic. Attention is one, albeit essential part in the ATCO’s task, important at any moment during task execution.

- In classical vigilance tasks the target signal is clearly defined, while in ATC different target signals with regard to their salience are present. Such a target might be a Short Term Conflict Alert (STCA) announcing an incident, the occurrence of a separation infringement itself or any type of deviation from the mental picture previously formed that potentially contributes to an incident. For the latter a continuous change in the constellation of elements and background composition is noted.

- Insofar as it concerns the complexity of the task, the amount and the nature of information presented to an ATCO are strongly varying over time (e.g. night shift is characterized by rare events) and involve auditory and visual information from multiple sources.

- The action cycle of controllers is not complete after signal detection but further evaluations and decisions have to be undertaken. ATCOs have to continuously monitor tasks and detect target signals, which are typical for vigilance tasks, but consequently they need to integrate the information and readapt their mental picture.

- In addition, the ATCO does not wait for warning signals, since he or she can avoid a critical event through active, preplanning behavior. Conversely, a state of readiness in vigilance tasks is comparable with passive behavior. This argument is similar to the vigilance definition of Hockey and Tattersall (1989) who discussed vigilance as a state in which the user is generally alert and actively involved in searching, problem-solving, predicting and planning.

In summary, as for many concepts also vigilance is a rather unclear and manifold approach despite its long history in research and different foci that have been set. Overall, compared to monotony, the vigilance concept is broader and leaves more space for ambiguous interpretations. At the same time, there is hardly any awareness about the different attentional concepts contained in the vigilance concept as discussed in the past. Finally, it is remarkable that, whatever conceptual description is used as a background for research, vigilance can be impaired in monotony, fatigue, or stress, even though for different reasons. For example, one cannot detect a signal because of an insufficient state due to fatigue or due to monotony. Even though it is agreed to see vigilance as a positive aspect required during task execution, a lack thereof demonstrates the need to distinguish various critical states concepts, which is not covered by definitions around low vigilance.

2.1.6. Stress

To maintain attention when there is little to do, is often considered as tiring and stressful (O'Hanlon, 1981). Increased strain in such conditions has been explained with a mismatch between the current operator states and the desired state (Hancock & Warm, 1989). The physiological changes when executing vigilance tasks were also described in relation to stress (Frankenhaeuser, 1971a). In contrast, Melton, Smith, McKenzie, Wicks, and Saldiver (1977) reported that ATCOs in low density air traffic control centers also had low stress levels. An overview about this early stress research in ATC summarizes these outcomes (Smith, 1980).

In a literature review, Thackray (1981b) examines the often postulated relationship between boredom/monotony and stress. After reviewing studies he concluded that the results do not support the hypothesis that feelings of monotony or boredom are accompanied by significant increases in commonly employed indices of stress reaction. Moreover, the total elements of the job have to be analyzed to find a connection to stress. The position is favored that stressfulness is appearing in monotonous, repetitive tasks in the case where the requirement for high alertness, continuous and rapid decisions and penalties for errors are coupled. It is noted that the discussions about increased or decreased arousal under boredom are linked to these arguments. Moreover, studies on monotony undertaken by Swedish workgroups (e.g., Melin, Lundberg, Derlund, & Granqvist, 1999; Lundberg & Johansson (2000), which are frequently used to support the stress argument in monotony, can be interpreted in this light (*cf.* section 2.4.).

Scerbo (2001) maintains that as long as one is required to work on a boring task, the task is stressful. This arises from the need to combat the boredom of having to continue working beyond one's satiation point. Hitchcock, Dember, Warm, Moroney, and See (1999) investigated if high workload is a consequence of the need for continuous signal observation (direct cost) or the effort to combat boredom (indirect cost). An experimental paradigm contained cueing of signals and knowledge of results with 108 students executing a vigilance task for 40 minutes. Finally, cueing resulted in a high-boredom/low-workload profile, which supported the direct cost model.

But any description of monitoring tasks as stressful needs to consider which understanding and framework of stress was used, e.g., Ognianova, Dalbokova, and Stanchev (1998) investigated alertness and sleepiness applying the term of stress states. A careful review of studies is necessary. In the cognitive transactional framework of Matthews (2001) stress arises when individuals appraise their environment as exceeding their resources. Desmond, Matthews and Bush (2001) concluded stressfulness as a consequence of simultaneous and successive vigilance tasks after 50 participants in a 48-minute-vigil-task filled in the Dundee Stress State Questionnaire (DSSQ). At the same time, changes in subjective mood, tiredness, tension, motivation and loss of concentration were found. Similarly, Warm (1993) concluded the stressful nature of a vigil task from participants who rated themselves as less attentive and more bored, irritated, strained and fatigued at the end of the task compared to the beginning. Given these results, frustration might have acted as an intervening factor through increased irritation. At this point the satiation concept comes into play. An alternative interpretation of the fact that Scerbo found high workload in vigilance tasks attributes the individual perception of high task load to continuous signal detection.

Thus, the stress concept applied around monotony is embedded in a very specific context and does not consider different approaches towards stress. A high number of definitions is available, leading Buunk to the statement that “there seems to be one aspect of stress where most researchers agree on, that is that there is a confusion of definitions of stress” (Buunk, 1998, p. 148). In a general view, stress models are distinguished that include stress as a cause, a reaction or a transaction/mediation. This recalls the definition problems already outlined in the section on monotony. Most contemporary researchers have accepted the transactional model of Lazarus (e.g., Lazarus & Folkman, 1984) where stress is a result of a transaction between a person and the environment. If a person appraised this as a threat, coping is started to handle the situation. There it may also be linked to temperamental traits, where coping is an individual style to avoid or reduce stress (Strelau, 2001). Threats are also connected with negative emotions, which contributed to a change in the focus of Lazarus’ work centering emotions as a crucial feature in the stress process (e.g., Lazarus, 1993).

Overall, it seems that the application of different stress concepts led to inconsistent conclusions. However, it cannot be ignored that at a certain point repetitive and uneventful work situations might swap from (objectively) underloading to (subjectively) overloading conditions. If stress is the consequence of a lack of resources related to coping strategies, this reaction occurs regardless of involved task characteristics. Ulich (2001) stated that monotony, vigilance, boredom and underload exceptionally result in stress but overload always leads to stress. As it was pointed out, stress is a possible consequence in the context of repetitive or uneventful tasks, but for different reasons than the ones proposed by Hancock and Warm (1989). In the context of this thesis, the underlying stress concept is used as proposed by Richter and Hacker (1998), who defined stress as a complex psychophysiological reaction to an experienced threat in work.

2.2. MONOTONY EMBEDDED IN THE PROCESS OF WORK

A variety of models and theories exists to explain how task execution affects an operator in a specific work environment. The following section describes the basic assumptions of the most essential ones and discusses their relation to the concept of monotony. It is remarkable that a different focus has been set if the approaches are compared internationally (Haga, Shinoda, & Kohubin, 2002). While North American research is dominated by assessing the mental workload when designing systems, Continental Europe’s ergonomists focus on the description of work consequences. The goal of work psychology is seen in the increase of production efficiency while guaranteeing protection from physical and psychical impairments and guaranteeing the development of the personality. European work psychologists do not only consider short-term effects but also long-term consequences of exposure to a work situation when addressing the optimization of personality and health promotion (Rau & Richter, 1996).

Apparently, the independent development of Occupational Health Psychology in the Anglo-American field allowed the establishment of human factors and engineering as a separate branch. As a consequence, only few models combine multiple temporal dimensions to describe the individual work consequences. An example is represented in Matthews and Zeidner (2004) who discussed the connection of adaptive short-term and long-term processes through the inclusion of personal goals and personality traits. In this context, self-regulatory processes take place on several levels to reach defined performance goals. In addition, the interaction of individual, task and environment as it was established by Russian work psychologists (e.g., Leontjew, 1982), relatively recently gained influence in the Anglo-American research. On a short-term basis, Manzey (1998) differentiates energetical models, resource theories or a combination of both to describe immediate work consequences. Mid-term effects, as they occur between two working days, are regulated through personal resources, recreation and recovery processes. They merge into long-term consequences concerning health, well-being and job satisfaction. This categorization is also maintained in the next section.

2.2.1. Concepts to Explain the Effects of Task Execution on the Operator

Workload is a commonly used expression to refer to the subjective experience of task difficulty. A clear distinction between the internal and external individual worlds helps to reduce the variety in understanding this term. This was already stretched when describing monotony as a cause and a consequence. The outside world is affecting the individual while the individual is reacting to his or her environment, and thus interacting with a task to fulfill the requirements. Finally, the perception of workload is not considered as negative per se unless deviations from an optimum range on the continuum between underload and overload occur.

The differentiation between taskload and workload for the domain of ATC was clearly outlined by Hilburn and Jorna (2001). They define task load as the demand imposed by the ATC task which consists of airspace factors (e.g., traffic load, number of traffic problems, flight altitude transitions, aircraft mix, and weather) or interface demands opposed to workload as the ATCO’s subjective experience of the demands. The link between taskload and workload is seen as a causal one mediated by skill, training, experience, and fatigue. Moreover, the experienced workload depends on the invested effort, employed strategies and observed performance (Tattersall & Hockey, 1990, p. 384).

This framework is closely related to the stress and strain concept (Luczak & Goebel, 2000) which explains that external task demands imposing *stress* on an individual result in psychophysiological reactions (strain) while fulfilling these demands. Originally applied to physical work, this concept was transferred to mental tasks in the early 1970s (Rohmert, 1973; Rohmert & Luczak, 1973; Luczak, 1975; Rohmert, 1984) and represents the basis for ISO 10075. Its basic assumption is that the entire external influences a human operator is exposed to will result in individually perceived strain. Depending on individual and/or actual conditions, facilitating (e.g., activation, warming-up) or impairing effects (e.g., mental fatigue, fatigue-like-states, and satiation) emerge. However, it needs to be noted that the development of this standard was strongly promoted by industrial needs (Nachreiner & Schultetus, 2002). Still, scientific support for the distinction of different critical states is insufficient and long-term consequences were neglected. Exceptions are the demand-control model predicting a relationship between jobs characterized by high demands and lack of control with increased stress and physical illness (Karasek & Theorell, 1990), or the investigation of relationships between short-term strains and burnout (Demerouti, Bakker, Nachreiner, & Ebbinghaus, 2002). Although criticized for its unsatisfactory theoretical status (Nachreiner & Schultetus, 2002), the standard provides a framework for classifying and integrating a variety of psychological phenomena in the work environment.

However, it is considered that there is no unified use of *stress* or *load* in translated publications from German researchers to characterize the objective task demands impinging on an operator. Depending on the underlying framework, they represent different components in the work process. For this reason, a discussion of these terms and similar expressions is included. Sanders (1998) pointed out that questions of load and stress are clearly connected with energetic aspects. Gaillard and Wientjes (1994) describe mental load and stress as related concepts, originating from different theoretical frameworks. In the German language *load* (*Belastung*) has a negative interpretation in the sense of burden, while *strain* has a negative connotation in English usage. But if translated from the German term *Beanspruchung*, the latter is meant to have a neutral meaning and can be either positive or negative. According to Greif (1991) the German term for load excludes inner triggers, while mental load includes inner conditions. Buunk (1998) equates *strain* with stress reaction. However, to avoid misunderstandings in the current work it is preferred to use (task) load for objectively measurable working conditions and strain or workload for the individual reactions, wherefrom critical states such as monotony, fatigue, or stress may emerge. As a final remark, the load-strain-model is to a certain extent related to a stimulus (load)-response (strain)-model. There, the operator represents a rather passive element. The execution of an activity is rather considered in the following theoretical framework.

Developed in the context of the activity theory (*cf.* Bedny & Karwowski, 2004, for a review of activity theory), action regulation theory (Hacker, 1986; Hacker, 2003, Richter & Hacker, 1998) explains the relationship between load and strain with a focus on active psychic regulation of actions according to task goals. The work process is described as a goal-oriented activity where actions are regulated by a hierarchy of goals and plans. An activity can be analyzed on three hierarchical levels, which are (1) the activity as a whole; (2) coordinating objective purposes and motives of the operator; and (3) distinct actions required to reach task goals and the specified operations to execute an action (Leonova, 2003). The sequential phase starts with *action preparation*, which means the orientation towards the task and its conditions, available methods and strategies, and degrees of freedom. This is followed by *action implementation*, a phase guided by continuous feedback on goal accomplishment and completed with an *evaluation* of the final outcome in terms of the task criteria. Applicable modes of control include automated, knowledge-based and strictly conscious intellectual models, which Hacker (2003) carefully differentiated from Rasmussen's levels of information processing. The hierarchical-sequential pattern of task execution has to be complete for a flexible and efficient action structure.

A key element of the concept is the redefinition of the task through emotional and cognitive evaluation of the task goals relative to own performance capabilities. Strains arise in the regulation of actions during active, goal-oriented coping with the task. Different work strategies, such as increases in effort, unspecific general activation, changes in work strategy and task goals mediate the coping process. This process cyclically switches between destabilization and efforts of restabilization and strains the performance capabilities of the individual (Richter & Hacker, 1998). In consequence, changes in the object of work as well as in the individual occur. Positive individual consequences comprise motivation, learning and personality development, while negative consequences of strain include not only fatigue, monotony, satiation, or stress, but also the loss of qualification. Zapf (1999) summarizes evidence for negative health effects in consequence of cumulating action regulation problems over time. Having an impact on the selection of activities and conditions to reach task goals, thus having control, is another essential component for successful task accomplishment. Zapf (1995) sees the hierarchy of goals and plans necessary to carry out a task as an approximation of complexity, while variety is expressed in the number of different actions required on the sequential dimension and thus independent of task complexity.

Even though this model allows building a gap between cognitions and actions in task execution, it does not sufficiently consider the cognitive processes required by the operator, a common characteristic shared with the load-strain-model. Nonetheless, Wieland-Eckelmann (1997, p. 431) sees attentional aspects, central in resource theories, implicitly discussed by Hacker who describes the regulation of signals from the working task as an essential component.

One of the most frequently used definitions for mental workload (MWL) refers to the amount of the operator's limited processing capacity needed to perform a given task (0'Donnel & Eggemeier, 1986). This statement is based on resource theories which date back to the 1970s. Sanders (1997) provides a summary of resource theories which generally postulate that a certain amount of resources is required to process a task. This capacity model of attention was proposed by Kahneman (1973) who defined capacity as the total amount of attention available for processing. An advancement of this theory, multiple resource theory (e.g., Wickens, 1992), posits that there are separate pools of resources along three dimensions that are defined as the involved modalities, the processing codes and the stage. Processing occurs on an early or late stage, verbally or spatially, and through visual or auditory input modalities. Several criticisms have been expressed concerning resource theory. One example is the problem of assessment since, hypothetically, a lack of resources is directly related to performance impairments (Matthews, 2000). Additional weaknesses are described by Szalma and Hancock (2002).

Despite these shortcomings, a further development of this theory was undertaken by Young and Stanton (2002a,b) to explain underload. The malleable attentional resources theory (MART) hypothesizes that there is no constant pool of attentional capacity, but resources may shrink to accommodate reduced demands, resulting in inefficient effort mobilization if performance is required. Their theory was supported in a driving simulator experiment (n=30 students) where vehicle automation was manipulated at four levels ranging from manual to fully automated driving. A secondary task and eye movement recordings were used to assess MWL. A decrease in MWL was found with increased levels of automation. The allocation of attention to the secondary task as reflected in eye movements and performance indicators became less efficient, rather representing shrinkage of resources than a change of strategy. In a further study Young and Glynick (2005) compared flight simulator performance in an underloading and normal condition. Ten participants had to maintain the same flight level when demands were low and continuously adjust altitude in the normal condition. The attendance towards a visual-spatial secondary task was used to measure spare attentional capacity through the number and mean reaction time of responses, also the reaction time to an additional critical event, sudden crosswind, was collected. That no significant results were found might be explained with the short time-on-run (10 minutes) and that performance decrements are not always visible, as operators use strategies to counteract impairments (Hockey, 2003). A multivariate assessment might have helped to identify if an inappropriate state contributed to the results. Additional shortcomings of this theory are that there is no clear definition of resources and the question remains open as to why operators do not actively accommodate their resources if increased workload is expected. Overall, in comparison with action regulation theory MART does not consider the active involvement of an operator. Also, the subjective evaluation of the process is neglected and results can also be interpreted as deactivation processes. However, in agreement with the conclusion of Wieland-Eckelmann (1997, p. 436), if attention shall be included as a basic cognitive function at work, the action regulation approach and the resource theory approach need to be combined. Gaillard's (2005) concept of concentration might contribute to link energetical and resource models. When concentrating, a person must continuously and purposefully regulate energy, function and precision of his/her actions.

This leads into a further discussion of the concept of effort that was originally introduced by Kahneman (1973) as the capacity or attention available to perform a task. Pribram and McGuinnes (1975) proposed a three-process neuropsychological model of attention, where effort was described as the coordinating process between the stimulus-determined arousal and the activation controlling response readiness, which is necessary to uncouple arousal and motivation. This was elaborated on in the energetic resource model of Sanders (1983) which combines structural/cognitive and energetical components to describe task performance. The structural level describes the flow of information through various processing stages from stimulus to response. At the energetical level the mechanisms described by Pribram and McGuinnes were distinguished.

As pointed out by Mulder (1986) arousal and activation are determined by involuntary factors, whereas effort is under voluntary control, mediates response selection and coordinates the arousal and activation subsystems. This aspect of voluntariness apparently distinguishes his understanding of effort from the one of Kahneman. When a suboptimal state of the organism has to be adapted, effort denotes the compensatory mechanisms involved in such a state-control process (Mulder, Mulder, Meijman, Veldman, & Van Roon, 2000, p. 150). In addition, the authors differentiate computational effort from compensatory effort. The former is related to processing complexity of tasks in the tradition of Kahneman (1973), whereas the latter is relevant if tasks are performed in adverse conditions. Kok (1997) pointed out that energetical mechanisms provide the gain for the data processing system and have direct ties not only with processing stages, but also with state variables like fatigue, drug effects and so on. The actual psychophysiological state of the organism may be suboptimal for meeting the demands placed on it and require an adaptation. Such state-regulating mechanisms are stressful in the conceptualization of Mulder and coworkers, which is characterized by elevated rates of adrenaline and cortisol and subjectively by feelings of tension and anxiety.

The aspect of self-regulation is emphasized in the state control model of Hockey (1997) that builds on Sanders (1983) idea of combining cognition and energetics. The goal of the model is to describe how task performance is adapted to the actual subjective and physiological state. He posits two hierarchical levels of control where the lower level describes normal, routine and skill based task performance and the high level control system regulates the instigation of effortful activities aiming to compensate for sub-optimal internal states. So, effort is located centrally as a coordinating process, adjusting the balance of input and output operations. Effortful regulation refers to the attempt to maintain a particular task state under overload, external distraction or stress. Already in precursory work for the model, Hockey (1986) has described the regulatory process of state control in which active direct or indirect coping strategies were distinguished from passive strategies, but also established a connection to some aspects of Hacker’s action regulation theory. Failures in regulation occur if one needs to maintain a vigilant state for a long period, in extreme environmental conditions, in suboptimal internal states like stress or fatigue and excessive workload. The prolonged active management of resources that are required to perform a task can lead to a deterioration of performance that has implications for short-term well-being and long-term health. This model explains the variety in psychophysiological reactions found when exposed to tasks.

Noteworthy are also the dissociated results in physiological and subjective measures as for example reported in Veltman and Jansen (2003). In an experiment with 11 pilots comparing the effectiveness of 2D and 3D radar displays in fighter aircraft, performance and subjective effort were significantly affected by the type of display, whereas the effect of workload was rather shown in physiological measures. The authors explained the results with the nature of the task that reflected data-limited aspects in subjective and physiological data, whereas resource-limitations were only obvious in physiological reactions. This led to the development of a framework where the required operator state was integrated as a crucial component in information processing as it is related to task goals (Veltman & Jansen, 2004). Postulating an information processing loop and a state regulation loop, required performance can be achieved through adjusting the state, adjusting the intensity of information processing or changing the task goals. The strength of this model is that it integrates relevant aspects from action regulation, state control theory and cognitive models and includes task goals, contexts and stressors.

The task goals are also central in the concept of the operator functional state (OFS) that is “the ability to carry out the job at that moment in time” (Wilson & Russel, 2003a). It refers to the multidimensional pattern of processes that mediate task performance under stress and high workload, in relation to task goals and their concomitant physiological and psychological costs (Hockey, 2003, p. 8ff.). Its analysis includes specific demands of the task and environmental conditions, the current acute and chronic operator conditions, the pattern of interaction with task goals and the stable operator characteristics like skill, motivation and coping style.

But as emphasized by Hockey, the behavior is not only constrained by task goals, but also other long- and short-term factors named somatic goals, i.e. motivational goals to satisfy basic needs and emotional goals to guarantee well-being. The modes of active response are engaged when performance is optimal in conditions of high demands and control. Under high demands and low control the operator can be strained where the performance is still adequate but the perceived level of demands is too great to be met within the working effort budget and thus primarily contributing to fatigue. Frankenhaeuser (1986) also described strain corresponding to the coping pattern of effort with distress. Disengagement occurs when high demands allow low control resulting in reduced performance. Hockey identified four patterns of latent decrement under stress and high workload where (1) performance can be impaired in the achievement of secondary goals; (2) the strategic adjustment makes use of less demanding cognitive operations; (3) the regulatory costs increase sympathetic activation and effort; and (4) low effort is used in post-tests.

In summary, the various models describing how the individual is involved in the work situation do focus on very different aspects. The facets that have been worked out indicate that several links to the concept of monotony are possible and present thus a potential for an overall integration. While they look at the description of individual states, task-related activities and cognitive processes, it is however not clear how they interrelate with monotony. Also, it seems that the models do hardly allow fully integrating the described phenomena related to monotony. Few of them do explicitly mention monotony, and they might also be different in their application for uneventful or repetitive conditions. However, most of the models do have certain elements relevant to the prediction of monotony.

2.2.2. The Work-Recovery-Cycle and Other Mediating/Moderating Factors

In the last section it became obvious that most of the models explained performance when directly related to the process of task execution. Comparatively few researchers integrated recovery processes and preexisting states or traits in their performance models. This is the focus of the next section.

Classical studies by Kraeplin (1903), ergonomic studies on physical work, the initial state approach of Kallus (1992) as well as the cyclic model of Wieland-Eckelmann and Baggen (1994) integrate recovery as a complementary process to strain to re-establish strained resources before encountering new tasks. The psychophysiological initial state precedes the strain process that is followed by critical states. Recovery is the return of physic and psychic indicators to a baseline level after task execution. Sluiter, Frings-Dresen, Meijman, and Van der Beek (2000) proposed to categorize this process in four time periods. Reactivity still occurs during the activity and needs to be considered because of the recovery effect of micro pauses (Meier, 1984). Meso-recovery is the period until one hour after the task, meta-recovery comprises one hour after work up to two days and macro-recovery happens on a long-term period after two days. To provide homeostasis between strain and recovery is one of the basic principles for work design (Luczak, 1998, p. 279). The human system has the tendency to maintain the balance of the body referred to as homeostasis[3]. Consequently, a balanced system provides the basis for optimal human processes and behaviors expressed in good performance.

For Meijman and Mulder (1998) physiological and psychological reactions during the work are an adaptive reaction to the working conditions and personal effort. In their model for effort and recovery they postulate that these reactions are reversible as the emotional, cognitive, and behavioral symptoms turn back after effort stops. Recovery means that when the imposing load decreases, the psychophysiological system is stabilizing to an initial level before the strain. But recovery is only efficient if there is enough time and possibilities to recuperate.

[3] In this context also the term allostasis (McEwen & Winfield, 2003) needs to be mentioned, which rather refers to the adaptation of the body to a dynamic balance.

When strain is accumulating over time without sufficient recovery to establish one's homeostasis, residual symptoms of the previous effort remain and also small stress reactions can show negative impairments. On a short-term basis sufficient recovery is essential, as otherwise individuals will start work with an impaired initial state, a subject investigated by Kallus (1991, 1992). Different techniques exist to support successful recovery. Typically one might think of the recuperation effect of rest breaks, leisure activities and work leave. Initially it was thought that regular, frequent rest breaks would be sufficient (Graf, 1961), but later on the importance of active and passive rest breaks, depending on the preceding type of strain, was empirically confirmed (Loehr & Preiser, 1974). The introduction of appropriate rest break cycles avoids the accumulation of strain. For example, Boucsein and Thum (1997) confirmed in patent examiners that short and more frequent rest breaks had a positive effect only in the morning while in the afternoon longer breaks were more effective. It also needs to be noted that a break consists of various processes and functions that may partly overlap, pictured in the floodgate function (Eberspaecher, Hermann, & Kallus, 1993). The wrap-up serves for regeneration, regulations and coping following the preceding strain, the passage phase and preparation phase reorient for the subsequent exposure to the task.

A theoretical assumption is that insufficient recovery has long-time impairments for health and well-being, e.g., in the manifestation of burnout, since the strain is cumulating over time (Meijman & Mulder, 1998). Elevated sympathoadrenal activation as a consequence of insufficient recovery after work was related to health problems in truck drivers (Kuiper, Van der Beek, & Meijman, 1998). To indicate early fatigue at work the concept of "need for recovery" was developed, which is characterized by temporary feelings of overload, irritability, social withdrawal, lack of energy for new effort and reduced performance (Van Veldhoven & Broersen, 2003). Its intermediate role between unfavorable work demands and subjective health complaints was confirmed in several studies reported by Sluiter, De Croon, Meijman, and Frings-Dresen (2003). However, the relationship between short-term strain and health, well-being, job satisfaction and absenteeism is not clear. Janssen, Kant, Swaen, Janssen, and Schroeder (2003) found in 7495 employees in different branches that the level of fatigue at work predicted the time of the first onset of sickness absence. The prevalence in three- or five-shift-schedules was confirmed in further analyses on 12095 employees (Jansen, Amelsvoort, Kristensen, & Van den Brandt, 2003).

A major mediating factor in strain processes is the existence of resources. Already included in resource theories to explain performance, resources can be approached from different sides. Hobfoll (1998, p. 45) defines resources as the objects, conditions, personal characteristics, and energies that are themselves either valued for survival, directly or indirectly, or that serve as a means of achieving these resources. Schoenpflug (1986) distinguishes external from internal personal resources. Within the latter, structural resources can be used without being depleted, such as the working memory or intellectual abilities, while consumptive resources are energetics that are consumed and need to be regenerated. Efficient behavior requires an ideal combination of both. In this light, the resource term as used in resource theory is only one specific form of consumptive or structural resources, a distinction often not considered by representatives of resource theory, even though indicated in the mentioned review of Szalma and Hancock (2002). But already in 1991 Schoenpflug emphasized that the interplay between these different types of resources is a way to bridge the gap between resource theory and activation theory under consideration of regulatory processes for the purpose of economic use of resources. In his opinion, arousal is the allocation of consumptive resources while the consequences of activation are the preparation of structural resources. Energetical resources mean basic support mechanisms as arousal, activation and effort to adapt individual psychophysiological resources to the task and it can also be seen as a general potential for behavior (Wieland-Eckelmann & Baggen, 1994, p. 113). Palliative resources have also been described in the stress research around Lazarus and mean the ability to regulate emotional states. As such stress is a thread of resources. Coping strategies may be seen as resources, because individuals dispose of various strategies to cope with reactions in an active or passive way.

In summary, the relevance of considering the development of strain processes in a long-term period has been demonstrated with different arguments. For an efficient evaluation of monotony in the field a thorough analyses of the surrounding factors is necessary. The early recognition of a potentially favoring situation to develop monotony would allow integrating countermeasures at this stage.

2.3. THE MULTIDIMENSIONAL ASSESSMENT OF MONOTONY

To study human behavior when executing a task, three basic measurement approaches are used, namely the collection of physiological, subjective and behavioral indicators summarized as psychophysiological assessment. Psychophysiological assessment evaluates if the human is in an optimal state to perform a particular task. Physiological measures help to gain information not accessible with behavioral measures and are considered to be objective. This is especially the case, as negative consequences such as fatigue or stress may remain masked for a long time. Thus, the so called multi-level-approach has the advantage of combining different levels of empirical data. Nevertheless, physiological and psychological perspectives are complementary, as dealing only with physiological events might lead to restricted descriptions or erroneous interpretations. A person perceives the task in a certain way and relates it to subjective emotions. Thus, the combination of objective and subjective measures is necessary to understand the variety in human behavior at work as it is also relevant for the occurrence of monotony. Additional aspects that need to be considered will be described in the following sections.

2.3.1. Basic Underlying Concepts

Arousal and activation have been seen as basic concepts underlying successful task execution since Yerkes and Dodson (1908) reported an inverted U-shaped curve relating the level of arousal to performance. However, these two terms have often been confused. While activation designates the tonic or long-term component of physiological activity, arousal refers to the phasic or short-term response. Hockey, Coles and Gaillard (1986) even introduced the term *energetics* to avoid misunderstandings around these concepts that express the intensity aspect of behavior.

Initially assumed to be a uni-dimensional concept (Duffy, 1957), support for multidimensional arousal was presented by several authors. Distinguishing cortical, autonomic and behavioral arousal, Lacey (1967) described manifold response patterns as *directional fractionation*. Kahneman (1973) suggested that arousal is not a passive process, but can be regulated by environmental and task demands. In the words of O'Hanlon (1981), arousal can be task-optimal with respect to performance and personal-optimal as determined by the homeostatic set-point. In subjective arousal Thayer (1967) distinguished tense arousal along the continuum from calmness to anxiety from energetic arousal that is ranging from tiredness to energy. Dickman (2002) empirically distinguished the dimensions of energetic arousal into wakefulness (sleepiness – alertness) and vigor (physical fatigue – readiness to engage in physical activity). In addition, in 141 participants completing a reading comprehension task, a curvilinear relationship was found during the course of the day when vigor was related to performance. Wakefulness showed a linear relationship connected to greater carefulness whereas increases in tenseness resulted in faster and less accurate performance, leading to discussions concerning the scheduling of cognitive tasks depending on the dominating form of arousal.

On a physiological basis arousal is controlled by the reticular formation located in the brain stem. (*cf.* Fisher, 1998, p. 43). This structure exerts excitatory influence on the whole brain through the ascending reticular activating system (ARAS) that contains also nonspecific afferents. While tonic changes are hypothesized to be mediated through the lower ARAS, phasic changes are mediated through the upper ARAS. Sensory inputs stimulate the ARAS and it enhances the cortex to allow conscious perception.

If this perception is relevant, arousal is conducted back to the ARAS, serving as a feed-back system while other structures are responsible for attenuation of the cortex. A close relationship between the vegetative nervous system and the activating and inhibiting structures regulate changes also in the inner organs, e. g. increases of heart rate and blood pressure or adrenaline (e.g., Robbins, 1997).

The ARAS appears to be important for the development of monotony as an appropriate level of stimulation is considered to relate to alertness/wakefulness, the tonic activation of the brain. A concept used to explain monotony is habituation that occurs if stimuli without direct significance are acting repeatedly on the nervous system. In consequence this leads to a decline in physiological reactions as reflected in decreased electrodermal activity. On the other hand an orienting reaction is the mobilization of resources in confrontation with a new or challenging stimulus that requires shifting attention. Amongst the physiological reactions are an increase in electrodermal activity, a decrease in phasic heart rate, constriction of peripheral blood vessels and alpha-blockade in EEG (Andreassi, 2000). It needs to be kept in mind that the above described concepts address phasic arousal. Tonic components are dominant in the assumption that the organism is adapting to the situation and shutting down its energy level when there is too low or too uniform stimulation to ensure homeostasis (note the similarity to MART). Apparently different physiological systems are involved in either assumption. Else it might have a different effect on the concerned physiological systems if activation is reduced due to inappropriate task demands or due to fatigue after prolonged time-on-task. However, a sudden increase in demands does not activate all the components of the system as fast as would be necessary for a proper task execution and thus be the reason for reported performance failures.

2.3.2. General Principles of Psychophysiological Recordings and Monotony

Many authors use the expression *psychophysiological* recordings to designate *physiological* indicators with the argument that they indicate psychological processes. This goes back to one of the first definitions of Stern (1964) who said that psychophysiology is research where the independent variable is the psychological one and the dependent variable is the physiological one. Similarly, Cacioppo, Tassinary, and Berntson (2001) defined psychophysiology as “the scientific study of social, psychological, and behavioral phenomena as related to and revealed through physiological principles and events in functional organisms” (p. 5). However, the appropriateness of applying such a definition equating *physiological* and *psychophysiological* variables is questionable in the field of human factors, as it *a priori* excludes the earlier mentioned subjective and behavioral measures by definition. Rather psychophysiological measures comprise a range of physiological, subjective and behavioral measures. This is argued with the fact that the general underlying assumption of physiological indicators representing psychological processes does not generally hold true. Even though the goal of psychophysiology is to describe physiological concomitants to psychological processes, it expresses general changes in the organism as expressed in the activation concept (Scheuch, 1986). It has already been recognized by Fahrenberg, Walschburger, Foerster, Myrtek, and Mueller (1979) through the proposition of eight main indicators for activation that a sufficient description includes subjective measures. Also, some indicators, such as heart rate (HR), are well known to reflect different physical and psychic processes (Stemmler, 2001). Furthermore, subjective evaluations give indications about how the individual interprets the situation. Finally, behavioral reactions such as impaired performance can be caused by a variety of factors that do not necessarily need to be only of a psychological nature (e.g., extreme environmental conditions) and in consequence need to be well distinguished from psychological processes.

In the context of this thesis the term psychophysiology implies a multi-method-approach to investigate effects of work, as already discussed in Greenfield and Sternbach (1972). Also Thayer and Friedman (2000, p. 60) point out that no single measure or aspect of responding can adequately represent a complex latent construct. This argument considers system approaches that stress the variety of interactive mechanisms leading to similar physiological and behavioral states. This requires the discussion of basic methodological issues. The traditional approach applying analysis of variances (ANOVA) does not capture the temporal process in physiological changes as it neglects delayed physiological responses and compensatory homeostatic processes. Else, multiple sources of variance from persons, occasions and variables need to be recognized (Cattell, 1966). Cacioppo and Tassinary (1990) identified a psychophysiological outcome, marker, concomitant and invariant when they conceptualized the taxonomy of psychophysiological relationships in terms of their specificity (e.g., one-to-one versus many-to-one) and their generality (e.g., situation/person specific versus cross-situational/cultural). A slightly different focus was set by Foerster, Myrtek, and Stemmler (1993) who described stimulus, individual and motivation specific reaction patterns. This is one of the consequences of the so-called covariation problem that illustrates the low correlation between physiological variables. Stemmler (2001, p. 21ff.) cites different reasons, wherefrom the simultaneous input of multiple intrinsic rhythms and subsystems is a major aspect. However, there is a huge variety in approaches to response specificity, which makes it difficult to compare results (Hinz, Seibt, Hueber, & Schreinicke, 2000). Besides, physiological indicators differ in their curve characteristics, that is, how strong physiological indicators are reacting depending on the activation level. For instance, heart rate is more sensitive if activation is high while peripheral arterial tone is more accurate under lower activation (Schandry, 1996, p. 59). To cope with these difficulties Stemmler recommends the definition of main variables, the assessment of multiple components or investigation of activation as a process in natural settings.

The discussion of curve characteristics leads to the problem that the targeted system is already active before an experiment and remains active. The principle of initial value (Andreassi, 2000, p. 396) states that a particular physiological response depends on the prestimulus level of the system being measured. However, the initial assumption of Wilder (1957) that the physiological response is smaller if the initial level is higher could not be maintained due to the inconsistencies found in certain measures. Nonetheless the determination of the baseline level is an important issue as it changes as a function of subject, location and measurement (Gratton, 2000, p. 919). Different procedures have been proposed to handle this issue (e.g., Kallus, 1992), which will be further discussed in the method section.

2.3.3. A Discussion of Psychophysiological Measures

The application of a particular technique depends on its particular measurement properties, which are first of all reliability (a measure is stable within and across tests) and validity (the measure reflects the workload as intended or something different). O'Donnell and Eggemeier (1986) further propose the classification of measures along their sensitivity, diagnosticity, primary-task intrusion, implementation requirements, and operator acceptance. Sensitivity is the capability of a technique to detect changes in the amount of the load imposed by the demands. Diagnosticity refers to the capability of a technique to discriminate the amount of workload imposed on different operator capacities of resources. But it needs to be considered that these properties also might be interdependent. Additional problems concern the technical equipment and potential error sources during recording (*cf.* Fahrenberg & Wientjes, 2000), which is a delicate matter in field settings.

Physiological indicators

Physiological measures have the advantage of enabling continuous quantitative assessment and minimizing the subjective effects of the investigator. However, the often unknown equipment may influence a person's natural behavior. The necessity of assessing a reaction profile rather than single measures was discussed earlier. Physiological indicators are classified within different systems. Due to the dominance of the nervous system in the regulation processes for the current work the distinction between peripheral and central nervous system (CNS) is applied. Alternatively, Boucsein (1991) differentiated physiological indicators according to their ability to reflect physical, mental and emotional workload. Also, a three-arousal-model was presented in Boucsein (1991) and Boucsein and Backs (2000) that builds on the work of Pribram and McGuinnes (1975) and Fowles (1980). This model distinguishes between three types of arousal and proposes that each type is related to certain physiological indicators. For example, phasic heart rate represents *affect* arousal, HRV characterizes *effort* and tonic HR changes describe *preparatory activation*.

Several summaries and reviews offer a detailed description of the involved measures. In the context of this work, just the most relevant aspects are mentioned (*cf.* Andreassi, 2000; Stern et al., 2001; Cacioppo et al., 2000) and measures are allocated depending on the dominance of the peripheral or central system. The peripheral system refers to nervous tissue outside the brain and spinal cord and is further divided into the somatic nervous system controlling muscular activities (e.g., eye activity) and autonomic nervous system (ANS) controlling visceral structures (e.g., cardiovascular system). The ANS is divided in the sympathetic and parasympathetic branches that are innervating an individual in the phase of energy mobilization or in the phase of rest and repair. Cardiovascular measures comprise the assessment of heart rate, blood pressure, finger pulse volume and respiration. HR and HRV are assumed both to be influenced by the sympathetic and parasympathetic activity of the ANS. While the parasympathetic branch allows a very fast adaptation of heart rate within a second, the sympathetic branch causes slow changes in heart rate within up to 15 seconds (Berntson et al., 1993). Heart rate is usually indicated as the amount of beats per minute but can also be presented as the inter-beat-interval (IBI) in milliseconds, which is known to have superior statistical characteristics (Heslegrave, Ogilvie, & Furedy, 1979). Descriptive and spectralanalytic procedures as defined by the Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology (1996) are generally used to analyze HRV. The advantage of descriptive measures such as the standard deviation of heart beat intervals in 5-minute-periods (SDNN) or the average SDNN in 24-hour-recordings (SDANN) is that they are easy to calculate. Spectral analytic procedures determine how much individual frequency bands contribute to the variances in different classification bands. Table 1 presents an overview of the commonly used frequency bands and how they are interrelated with some descriptive indicators. Spectral analytic methods are more complex in calculation, but better differentiating as discussed in the literature (Berntson, Bigger, Eckberg, et al. 1997).

Table 1: The relationship between descriptive and spectral analytic heart rate variability indicators

Frequency Range (Hz)	Task force Band Name	Common Band Name	Function	Related descriptive
<0.03	Ultra low (ULF)	Low frequency	Circadian rhythm	SDANN[a]
0.03 – 0.05	Very low (VLV)		temperature regulations	
0.05 – 0.15	Low frequency (LV)	Mid-frequency 0.1 Hz component	connected with resonance of vasomotor system and blood pressure regulation	r-MSSD[b]
0.15- 0.4	High frequency (HF)	High-frequency	indicates vagal influence[c] and contains influences from respiratory system;	SDNN[d]

Note. [a]Average SDNN in 24-hour-recordings
[b]Square root of the mean of sum of squares of differences between successive R-R wave intervals
[c]Vagal tone is an index of the tonus of the vagus nerve, the main nerve of the parasympathetic branch
[d]Standard deviation of heart beat intervals in 5-minute-periods

Due to its easy application, heart rate and its derivatives have often been used in aviation (e.g., Jorna, 1993). Boucsein and Backs (2000) provide a systematic summary of physiological reaction patterns in laboratory and field studies. Increased heart rate has been mainly associated with increased strain and mental effort, while decreases were observed under fatigue, monotony, and/or low arousal. In air traffic control, HR was correlated with the number of aircraft under control (Costa, 1993; Rose & Fogg, 1993), and HR decreased during night shift (Costa, 1993). Brooking, Wilson, and Swain (1996) did not find an effect of task difficulty on heart rate in a lab study investigating eight ATCOs under low, medium and high traffic load and two types of complexity. This result can be explained with the concept of autonomic mode. Autonomic activity may change greatly during task performance but not be apparent in HR because of the modes of autonomic control (Backs, Lenneman, & Sicard, 1999; Backs, 2001), where sympathetic and parasympathetic activity might operate in a coupled or uncoupled vs. reciprocal or nonreciprocal mode. This was confirmed in a study with 16 female students who executed the ATC task in a similar setup as Brookings and coworkers (Backs, Navidzadeh, & Xu, 2000). The authors found that the mode of control differed across scenarios, as medium and high workload elicit reciprocally coupled sympathetic activation and parasympathetic withdrawal, were no change to baseline was found in low workload scenarios.

Mulder, Mulder, Meijman, Veldman, and Van Roon (2002) summarize findings of research on HRV. Generally, the execution of effortful mental tasks is accompanied by a decrease in heart rate variability or a reduction in the 0.10 Hz component. In exceedingly difficult tasks, individuals may choose to allocate less effort and consequently render the suppression of the 0.10 Hz component less obvious. This component was diagnostic in attention-demanding (resource-limited) control operations, but unaffected in data-limited processing. Also, in a study of Byrne (1993, quoted in Byrne & Parasuraman, 1996) individual differences were reported depending on how participants approached a task. High levels of effort showed a suppression of HRV and a faster reaction time while subjects with low effort showed an ongoing increase of HRV.

However, a systematic analysis of the measurement properties revealed a different picture concerning the validity of this component (see also Manzey, 1998). Nickel, Nachreiner, Zdobych, and Yanagibori (1998) found unexpected spontaneous fluctuations in this component that cannot be explained by theory. Else, HRV varies for different situation specific demands. Especially in field settings the use of this component is not recommended because of the quickly changing variations of factors influencing HRV. Further investigations were done on sensitivity and diagnosticity (Nickel & Nachreiner, 2003). In 14 participants executing a test battery sufficient sensitivity was only obtained if work and rest breaks were distinguished. As a conclusion the authors stated that HRV is only indicating emotional involvement and time-pressure as expressed in the discrimination between paced and unpaced tasks.

On the other hand, sufficiently high reliability was reported after a longitudinal and cross-sectional study conducted by Wagner, Rudolf, and Noack (1998) who investigated driving difficulty with HRV. Brookings et al. (1996) did not find any effect of traffic load or complexity on the mid-frequency component. Interesting as well is the application of blood glucose as a measure of mental effort (Fairclough & Houston, 2003). In contrast with cardiovascular indicators it was more sensitive to task load and time-on-task, while cardiovascular indicators were only sensitive to the latter. Overall, the results support the conjoint measurement of HR and HRV indicators together with the subjective perception of the situation when assessing monotony.

Systolic and diastolic blood pressures (SBP and DBP) refer to the force exerted by the blood against the walls of the blood vessels and drive the output of the heart through the circulatory system. As the regulated values, they are very influential in the cardiovascular system. An increase in systolic and diastolic blood pressure has been found with increased traffic load and more complicated tasks, but a decrease was associated with fatigue (Boucsein & Backs, 2000). In ATC, increased work stress and SBP reactivity have been associated with long-term risk of hypertension in ATCOs (Ming et al., 2004). It is however well known and needs to be considered in the interpretation of the data that lower heart rate can be interpreted as less activation or as increased baroreflex output because of high BP, which reduced the value of HRV as an indicator for mental workload (*cf.* Veltman & Gaillard, 1998). Therefore, Mulder and his colleagues proposed baroreflex sensitivity, which combines both HRV and blood pressure variation, as an index for mental workload (e.g., Mulder, Van Roon, Veldman, et al., 2003).

Respiration can attenuate or amplify the effect of mental load and thus needs to be considered in the measurement. Since respiration frequency ranges between 0.15 and 0.4 Hz it directly influences HRV (Sammer, 1998). Brookings et al. (1996) reported that respiration rate was higher with increased difficulty while respiration amplitude was not affected. Ohsuga, Shimono and Genno (2001) compared subjective and physiological effects of a monotonous tracking task with a stressful task and interpreted respiration instability as a reflection of boredom and disgust.

Electrodermal activity (EDA) describes the activity of sweat glands and is almost entirely innervated through the sympathetic nervous system. At the same time, circulating adrenaline has no strong effect. This system is specialized in mobilizing energy resources in response to internal and external milieu demands and responds to emergency situations. Nonetheless, diurnal variations (*cf.* Hot, Naveteur, Leconte, & Sequeira, 1999) in skin conductance level (SCL) and deactivation need to be considered (Boucsein, 1988). Nonspecific spontaneous fluctuations and increasing SCL were associated with higher activation (Boucsein & Thum, 1997) and the emotional content of the task (Schaefer, 1986). Averty, Athenes, Collet, and Dittmar (2002) reported SCL and heart rate as the measures that best correlated with task load index, aircraft count and subjective workload. Skin temperature is another measure reflecting sympathetic activity and often used in shift work to display circadian variation (Costa, 1999).

The perception and processing of visual information is a basic process in task execution where the EOG is collected through the application of electrodes around the eye. It is generated by a dipole formed by the cornea and retina and provides a variety of indicators derived from eye blinks and movements. Veltman and Gaillard (1996) discussed that eye blinks were specifically affected by visual demands and not by workload in general. This was confirmed by Van Orden, Limbert, Makeig, and Jung (2001) who conducted a study with eleven participants completing four 2-hour-blocks of an air warfare task. They found that blink frequency, fixation frequency, and saccadic extent were related to higher visual processing load. The function of peripheral perception is to drive attention towards a goal that needs to be fixated. The point of gaze is used to find out in which order the necessary information is processed. Many studies assume that fixation duration provides information about the mental load imposed by the task as it would require more central processing. However, David (2000) did not report significant differences in duration and number of fixations in simulated ATC depending on traffic load (15 vs. 30 aircraft in 30 min) or control mode (graphical mode or keyboard). Brookings et al. (1996) reported higher blink rate under low load, increased rates with decreasing complexity, but no difference in saccade measures.

Contradictory results for the impact of the task difficulty on eye blinks can be explained by a varying research paradigm (*cf.* Tanaka & Yamaka, 1993). Eye blink characteristics are also associated with fatigue (Stern, Boyer, & Schroeder, 1994, 1996). In two experiments with a total of 81 students calculating aircraft threat values Boehm-Davies, Gray and Schoelles (2000) confirmed that eye blinks are suppressed during high cognitive processing and increased after its completion. On the other hand, blink closure duration was better reflecting time-on-task effects together with flurries of blink (Stern, Boyer, Schroeder, Touchstone, & Stoliarov, 1994).

Saccadic movements describe fast jumps of the eye from one fixation point to another with a high number of saccades showing the necessity of visual searching processes. Moreover, they are more easily triggered under high arousal and as latency reduces. Saccade frequency has shown a significant decline during the course of an air traffic control simulation lasting for two hours (Stern, et al., 1994, 1996.). There is no clear indication of a decrease in saccade velocity, as it might be a consequence of increasing blinks (McGregor & Stern, 1996). With sleep deprivation, latency of saccades is increasing. Russo, Thomas, Thorne, Sing, and Redmond (2003) found that saccadic velocity is sensitive to sleep deprivation and sleepiness. A decrement in saccade velocity with increasing TOT was reported by Galley (1993) in drivers. App and Debus (1998) confirmed in two experiments with a continuous visual motor task (n=10 vs. n=16) a progressive decrease of peak saccadic velocity while performance was maintained on a high level. Target saccades rather reflect the high level of effort, return saccades demonstrate the basic activation level and therefore demonstrate how demanding the task is (*cf.* Galley, 1998).

Measurements of CNS activity consider tonic or phasic components of the electrical brain activity. The encephalogram (EEG) is supposed to reflect the cognitive demands placed upon an operator (e.g., processing information, making decisions, and initiating actions) as well as general alertness. Different EEG indicators are the sum of activity generated at several disconnected sites (Andreassi, 2000) and can be described through a decomposition in frequency bands, generally divided into beta (13-30 Hz), alpha (8-13 Hz), theta (4-7 Hz) and delta (0.5-3.5 Hz). Changes in spontaneous EEG activity were observed with changing mental demands, where a decrease in alpha and an increase in beta and theta were related with high mental load. Brookings et al. (1996) found increased beta-power in complex traffic situations and explained that the use of a different cognitive strategy may explain the lacking impact of traffic load. Alpha band power changes were driven primarily by the interactions between difficulty and traffic manipulations. EEG activity in the theta band was sensitive to task difficulty manipulations with increased percentage of theta power at central, parietal, right hemisphere frontal and temporal sites. However, the relationship between theta-activity and mental state is not clear. Pennekamp, Bosel, Mecklinger, and Orr (1994) found increased theta power in tasks requiring more attention and memory load. Belyavin and Wright (1987) reported rising theta activity with increased time on task. Yamamoto, Matsuoka, and Ishikawa (1989) found a positive correlation between frontal midline theta waves and performance in a highly concentration-demanding visual display task (VDT), further supported by Yamada (1998). Brookings and colleagues discuss that increased theta activity reflects different cognitive activity mechanisms, which are complex task performing and the arousal state. Recently, also the combination of different power bands in indices was reported as successful, for example the task engagement index of Mikulka, Scerbo, and Freeman (2002).

The relation to alertness was investigated, as reduced alertness is connected with higher power in alpha and theta bands. Makeig and Inlow (1993) pointed out that averaging data in groups will not yield as much information as studying individual performance especially when the individuals already differ significantly in their baseline levels. Belyavin and Wright (1987) investigated the relationship between EEG and vigilance in 9 subjects performing a simple visual vigilance and a letter discrimination task in seven sessions during 15 hours. Analyzing 3-second-periods preceding each critical stimulus, they found increased activity in the delta and theta bands, and decreased beta activity. The deterioration in performance was attributed to the decrease in arousal. They concluded that vigilance could be predicted more reliably from EEG than performance which involves relatively few errors.

Nonetheless, the prediction of vigilance levels from EEG is seen as problematic, since, in their study, it only differentiated epochs of very high error rates. In the assessment of EEG a variety of indicators needs to be considered. The most important ones are described in the guidelines of the American Clinical Neurophysiological Society. For example, diurnal variations have been found to be critical influences in vigilance tasks (Higuchi, Liu, Yuasa, Maeda, & Motohashi, 2001). Event-related potentials (ERP) are phasic indicators that have been discussed to reflect cognitive processing and mental load (Groß & Metz, 1998; Hohnsbein, Falkenstein, & Hoormann, 1995, 1998; Kok, 1997). In ATC, ERPs reflected a fatigue-effect as the relative amplitude of the P300 decreased significantly after high load exercises (Mollard, Cabon, & Bourgeious-Bougrine, 2000).

Finally, the excretion of hormones is under control of the ANS and CNS. Two basic systems, the sympathetic-adreno-medullary-system (SAM) and the hypothalamus-pituitary-adrenal-system (HPA), are connected through complex feedback loops within the neuroendocrine system. HPA activity is shown through the excretion of cortisol and increased under chronic stress. SAM activity is measured by concentrations of peripheral catecholamines and indicates activation reflecting the workload. This system is involved in situations with demanding active involvement or when confronted with challenging events. Increased sympathetic activity is responsible for the secretion of catecholamine in peripheral blood. Elevated noradrenaline levels characterize the mobilization of physical resources; elevated adrenaline levels are observed when mental resources are mobilized (Mulder, et al., 2000; Lundberg & Johansson, 2000). If the level of adrenaline and noradrenaline in the blood is elevated, it reduces hypothalamic activity, closing the negative feedback loop between hypothalamic activity, sympathetic neural output from the rostral ventral medulla (RVM) and adrenal medulla activity. The adrenal medulla activity maintains the level of sympathetic activation during long periods of task performance. The HPA-axis consists of hypothalamus, pituitary and adrenal cortex and is involved in stress reactions of long duration (hours, days, and longer), even though it shows differential reactions in laboratory studies with task durations of around ten minutes. If cortisol level is too low, CRH (corticotropine releasing hormone) is released that activates the anterior pituitary to produce ACTH (adrenocorticotropic hormone) to be transported by the blood in the vessels to the adrenal cortex. This results in production of cortisol in blood. Challenging situations demand an active mental involvement that provokes the adrenaline reaction, which becomes less pronounced and fades away, accompanied by feelings of activation. If an inadequate state has to be compensated in confrontation with situational demands, the elevation of peripheral adrenaline rates is embedded in a stressful syndrome characterized by feelings of tension and anxiety. Adrenaline is seen as an unspecific indicator for mental demand regardless of the effect. If there is a negative effect, an increase of cortisol is noted (Gaillard & Wientjes, 1994). The relationship of these systems helps to understand the interaction between mid-term effects of work and recovery processes.

Accelerated and long-lasting adrenaline excretion is known as a stress indicator that can lead to damaging health problems. In 22 ATCOs Melton, McKenzie, Polis, Funkhouser, and Iampietro (1971) found a relationship between catecholamine levels and the number of aircraft operations. Cortisol excretion showed an increase in the late night shift and incomplete recovery during the off-duty rest period. Rose and Fogg (1993) confirmed increased BP, HR and cortisol with increased aircraft under control but also defined different responders. In a sample of 381 ATCOs 20-25 % were high responders, but also an inverse response in HR and cortisol was described. Sluiter, Frings-Dresen, Meijman, and Van der Beek (2000) systematically reviewed research on hormone systems and their connection to recovery and reported that spillover often takes place after the work was completed. Incomplete recovery was shown in catecholamine and cortisol for different mental and physical tasks and time lags.

After all, there are some more aspects to be considered in the assessment of monotony. As physiological measures were mainly used to distinguish levels of task load or fatigue, Byrne and Parasuraman (1996) discussed their interpretation in the assessment of underload. The risk is to confound motivation and effort with task effects. Since task characteristics that elicit effort may not be dominant, physiological profiles depend on variations in compensatory effort.

Also, the importance of a baseline needs to be considered here, as recovery of a system is usually achieved when the physiological indicators return to baseline levels. Usually expressed through a decrease in indicators, the consideration of establishing homeostasis seems to be relevant in this context and requires a different description of the baseline. A further problem is the transfer of results from laboratory to field settings (Seibt, Boucsein, & Scheuch, 1998). A comparison of the reactivity of 58 sports students, who were exposed to lab tests and comparable field follow ups, revealed that heart rate was the only indicator allowing predictability (Fahrenberg, Foerster, Schneider, Mueller, & Myrtek, 1986). Suggestions were proposed to assess individual differences in the criterion situations themselves. Jain, Schmidt, Johnston, Brabant, and Von zur Muehlen (1998) observed that participants reacting stronger in the laboratory showed higher variability in the product of HR and SDP (RPP) and higher responses to stressful conditions in the field in cardiovascular and endocrine indicators. To conclude, meeting the requirements presented in the beginning of this chapter is essential. This is not only a problem of physiological measures, but relevant for any measures, as the thesis concentrates on the effects of conditions on monotony.

Subjective indicators

The subjective methods include interviews, questionnaires, and ratings or rankings. According to Schoenpflug (1987) data obtained with questionnaires reflect individual perceptions, which are the most important factor in mediating the stressor-strain processes. Often they focus on the perceived difficulty level but do as well address workload, effort, motivation, fatigue, or affective components. Amongst the latter, emotions or mood states need to be distinguished even though lay persons often use them synonymously. Davidson (1994) proposes a functional distinction, where the short-lived emotions prepare body and mind for appropriate immediate response and thus have the potential to bias action. On the other hand mood tends to bias cognitive strategies and processing over a longer term. The assessment of emotions at work has generally been neglected for a long time. This is surprising, as emotions come into cognition very early, even before information processing, as demonstrated in experiments of Oehman (1988). Autonomic reactions to unseen pictures were found to also lead to slower reaction time if considered fear-relevant. In this context the connection of the amygdale system, which is relevant in emotional processes, with the attentional system, is remarkable. Gendolla and Kruesken (2001) also confirmed the impact of moods on effort-related autonomic reactivity in active coping with a task.

Rating scales are frequently used to assess work-related aspects. NASA-TLX (Hart & Staveland, 1988, reprinted in Moray, 2004; Collet, Averty, Delhomme, Dittmar, & Vernet-Maury, 2003), Rating Scale for Mental Effort (RSME; Hilburn, Bakker, Pekela, & Parasuraman, 1997), Instantaneous Self-Assessment (ISA; Tattersall & Foord, 1996) and Subjective Workload Assessment Technique (SWAT; Reid & Nygren, 1988) are amongst the most applied in ATC to assess workload. However, various critics need to be considered in their application. Annett (2002) discussed that the judgment of humans only allows ordinal scale properties. Schuette (2002) confirmed diagnosticity and sensitivity of the effort scale only marginally. Additional arguments about the subjectivity and validity led to a series of discussions in edition 45 of the journal *Ergonomics.* Amongst others, authors addressed the belief systems, backgrounds of researchers, or contextual influences on definitions in the comparison of objective and subjective measures when results are contradictory (Young and Stanton, 2002c; McKenna, 2000). Salvendy (2002) believed that the reliability of the measures was critical, as any discussion on subjectivity or objectivity is reflected. He suggests framing subjective ratings in the context of psychophysics to significantly increase their robustness of measures.

Eilers, Nachreiner, and Boening (1989) tested the validity of unspecific one-dimensional scales to assess mental strain and concluded impaired validity in repeated-measurement designs due to reactivity. The judgment is influenced by intra-individual variations that occur because subjects do not use a fixed frame of reference but adapt it depending on the intensity and duration of a task. It is possible that rated task difficulty is just covarying with mental strain. Else, subjective expectations about research hypotheses can be generated. If using only rating scales it is not clear if operators work hard or think that they have to work hard.

If not asked for, individuals do not distinguish between demands by the task and the effort invested in the task. Gopher and Donchin (1986, quoted in Eilers et al., 1986) discuss that one-dimensional scales are dependent on a cognitive analysis of what is known of the task. Impairments on a task can just be measured when it was consciously reflected. Else, social context, normative conditions and resources affect ratings. Another issue was brought up by Schuette (1999), who reported an experiment where the repeated application of a subjective scale had comparable effects to short rest breaks. He recommended to avoid repeated-measurement designs and to use pre-post-test-designs especially when the rating follows a long period with low demands. Similar problems can be considered in multidimensional scales; however, they are usually developed applying a concrete theoretical background (e.g., DSSQ, Matthews, Joyner, Gilliland, Campbell, Huggins, & Falconer, 1999).

In comparison to subjective measures, Veltman and Gaillard (1998) demonstrated that physiological indicators were sensitive to mental effort, while subjective measures were sensitive to both, task difficulty and mental effort. Overall, their success is due to the advantages that they are easy to administer and commonly accepted by participants due to high face validity. But one needs to bear in mind the earlier described implications in the assessment of monotony. Even though critics are not of one voice, they need to be considered in the interpretation of results.

Behavioral indicators

In work settings, human behavior shows high variety when pursuing the goal to provide optimal performance. In ATC the occurrence of an incident or loss of separation is the most direct indicator of suboptimal performance. Nonetheless, a lot of errors may precede such a situation without necessarily resulting in an incident but a potentially critical situation. Therefore, the discussion of this section does not only include direct and indirect performance measures, but as well observable activities that announce or accompany suboptimal behavior.

Performance indicators directly reveal the quality of task accomplishment. Commonly applied measures in lab studies are error rates and reaction time as well as their variation. There are various restrictions in applying such measures. It has been observed that primary performance indicators do not necessarily reflect impaired performance. Therefore, secondary tasks were introduced to assess the engagement of a person in the main task. Unfortunately they have the disadvantage of potentially interfering with the primary task and making the task more interesting. Apart from that, interpretations of impairments are more complex as they do not only reflect the difficulty level, but also depend on influences like fatigue, motivation or learning effects. Hockey (2003, p. 12) summarized that performance decrements are generally small, selective and affecting less critical aspects of the task, and more likely to be observed in laboratory studies than in real-life environments. Prolonged work and stress are often necessary precursors while higher activation and effort counteract impaired performance. Various possibilities were identified for how performance can be affected. Some effects may appear only after the work period ended (Hockey, 2003, p. 18) and represent a sign that performance has been affected by a central fatigue state rather then a localized problem. Schellekens, Sijtsma, Vegter, and Meijman (2000) confirmed fatigue-after effects of mentally-demanding tasks executed on two simulated work-days (n=16) of low or high task difficulty. It was found that errors increased from the pre- to the post-test after the difficult day. Additionally, it was also shown in a delayed probe task after two hours that subjects shifted towards a low-effort/more-risk strategy indicated in HRV and error rates while feeling more fatigued and deactivated. However, the search for a sensitive test and carry-over effects can impose problems.

Veltman and Gaillard (1996) point out that the level of performance only provides valuable information when techniques are used to index the invested effort at the same time. The reason is that operators adapt to increasing task demands by executing additional effort. Support is provided by three studies of Galley (1998) with a total of 293 subjects. He found that performance was rather independent of a decline in activation whilst saccadic velocity was reflecting deactivation.

This led him to the description of a two-stage regulation process where, in a first stage, performance is stabilized by *concentration* that resists changing activation (e.g., through fatigue). When concentration fails to reach the goals, the *effort loop* becomes active through additional brainstem activation, which is reflected in saccadic velocity. In ATC, performance may be maintained through a change in the strategy. Sperandio (1971) described how ATCOs adapt their strategies to manage increased workload. Radio messages were more efficient and strategies changed from an individual to a more standardized strategy with increased number of aircraft under control. The change towards a less efficient strategy was also used as an explanation by Desmond and Hoyes (1996) who found that participants failed to mobilize their efforts effectively under low workload. Smolensky (1990) observed higher memory errors in situations in which an unexpected and rapidly downward-shifting of taskload level occurred and explained it with a disruption of the mental model. Relevant is the observation of Van der Hulst, Meijman, and Rothengatter (2001) in 24 participants that the safety margins were increased in fatigued simulator driving. While performance deteriorated in steering, no impairment was found in the unpredictable deceleration situation of high-priority. A further description of performance measures for ATC simulations can be found in Manning (2000). Also, a relationship between sector characteristics and operational errors was confirmed (Rodgers, Mogford, & Mogford, 1998).

As a commonly measured variable in classical vigilance task errors have been described for a long time. Bartenwerfer explained errors in monotonous situations as a result of the reduced ability of a person to react. Further impacts like boredom lead to coping behavior, which attracts attention from primary tasks. Edkins and Pollock (1997) investigated Australian Rail Accidents and found evidence of a high involvement of skill-based behavior. They discussed that slips and lapses can also occur during routine action sequences, which still require attentional checks while at the same time the full attentional energy is not reserved (Reason, 1992). Haider (1956) reported in 59 female industry workers executing an additional vigilance task that the number of non-detected signals and reaction time increased.

Not yet frequently considered in ATC is the factor of contextual performance (Touzé, 2005), which refers to taking over additional functions in the job. However, it might be interesting concerning the variety aspect of monotony, as additional engagement in work-related projects are a source of variation and job satisfaction for ATCOs.

Behavioral activities have been demonstrated to announce a change in the state. Rogé, Pebayle, and Muzet (2001) found a relationship between behavioral activities and performance impairment in vigilance tasks amongst 17 participants driving in a simulator for two hours. Non-verbal activities such as sighs and yawning precede physiological signs in the EEG indicators. Postural adjustments were more frequent during and after an increase in the low vigilance indicator while self-centered gestures occurred only after. It was discussed that these activities do have the function to either announce a suboptimal state or to reactivate a state. Increased motor activity under work monotony was also reported by Rzepa (1984).

Another approach used to predict workload of ATCOs is the analysis of communication events. Manning, Mills, Fox, Pfleiderer, and Mogilka (2002) analyzed the content of communication and found that taskload and subjective workload correlate positively with the number and total duration of communication events and negatively with the average duration of one event. However, the duration of communication is critical, as it may not only be used to pass information and thus correlate with workload (Porterfield, 1997), but it might also be indicating boredom as it is a chance to pass time in low workload situations.

To summarize, it is only through the simultaneous application of a number of measurement techniques that the effects of a particular work condition can be adequately evaluated. On a physiological basis, multiple indicators stand for the demonstration of the concept of multidimensionality. There is evidence for dissociation between subjective reports and other measures (e.g., O'Donnell & Eggemeier, 1986), which Brookhuis and De Waard (2002) see as an opportunity to better understand discrepancy between people's ratings and their behaviors.

However, only a combined approach allows the distinction of different critical states. This is also relevant for the assessment of monotony. Psychophysiological strain analysis must be interpreted together with performance if task demands provide the individual with the opportunity for operation management. The strain level can be kept constant to keep total efficiency high. For this reason, strain measures may mainly reflect ability for process optimization with respect to conditions of performance. Finally, a highly relevant aspect is discussed by Nachreiner (2002). He points out that it is often not clear if the intention is to differentiate between people according to their strain or to distinguish between the workload imposed by working conditions.

2.4. EMPIRICAL EVIDENCE FOR MONOTONY

This section provides an overview of the most important empirical findings on the issue of monotony. Due to the confusion surrounding the different terms for similar phenomena, studies on boredom, underload and vigilance were reported if they investigated the effects of repetitive or uniform tasks. They are classified as long as they focus on task characteristics (section 2.4.1), the embedment in an organizational context (section 2.4.2) or individual characteristics (section 2.4.3), even though different aspects may not be independent of each other. Unless necessary for the understanding, studies are reported that predominantly focus on cognitive rather than on physical tasks, whose nature is more similar to ATC than to assembly line work. This is underlined by the argument that ATC is a complex task with a high proportion of controlling and monitoring.

2.4.1. Task Factors

Following Johansson's (1989) distinction of uneventful and repetitive monotony, different task characteristics need to be distinguished. The early interest in monotony was a consequence of reduced performance observed in work conditions related to boredom and fatigue (Wyatt, Fraser, & Stock, 1929). Flechtner (1937, quoted in Gubser, 1968) reported longer task execution time and increased reaction time towards unexpected signals when performing a uniform task. A lot of basic work to understand the development of monotony was executed in the controlled conditions of the laboratory environment. Frequently, subjects were asked to repeatedly execute simple actions or monitor radars. Also Bartenwerfer (1957) developed his theory of monotony after studying the behavior of people in different experiments. In his main study 39 participants performed a driving task for 104 minutes under manipulation of speed, room temperature and time-of-day. Physiological deactivation in HR and HRV, reduced performance but increased variation in performance as well as subjectively perceived fatigue and sleepiness were found during task execution. Subjects were yawning, deeply breathing and closing their eyes. After task completion they stated the occurrence of startling along with the feeling of having been away for a moment or dreaming. In the early phase distracting thoughts were reported, while with continuous time on task subjects also noticed insufficient performance. The observation that performance was not impaired after changing the task led Bartenwerfer to the conclusion that monotony should not be equated with fatigue. In an additional experiment with 20 participants he compared the effects of a cognitively demanding task with a uniform (writing the number "8") and a relaxing condition. While the first one resulted in a state of fatigue, the second one resulted in monotony. However, the design of these experiments did not contain a control condition for the task variation. The connection to monotony was assumed through a description of task characteristics, but not through their manipulation. Arguments provided by Barmack (1939c, p. 470) support the distinction of fatigue and monotony as distinctive states. After adding numbers for four hours, participants verbalized their experiences predominantly as boring and not as fatiguing.

Generally, the task characteristics can be analyzed according to the frequency of stimuli that result in a reaction or require an action. The signal features received high attention in the context of vigilance research. After reviewing research, Wickens and Hollands (2000) identified signal frequency, signal display in terms of color or volume, time-on-task, and arousal level as the key factors in signal detection performance. Grandjean (1991) reported a finding of Schmidtke that the frequency of correctly detected signals increases with a presentation rate of between 100 and 300 per hour. Pfendl (1985) was interested in the effect of increased monotonous signals on performance. In the comparison of a serial and a parallel information processing task, no effect on signal detection was found, but monitoring performance reduced more rapidly under serial presentation. Hacker (1982) focused on the cognitive components in mental routine tasks. In 60 students the enlargement of arithmetic and logical operations in more complex tasks resulted in better performance, less errors, lower perceived fatigue, performance impairments, motivation, and boredom. In contrast, the enlargement of operative memory requirements resulted in the opposite outcomes. The effect that increased task demands are able to rule out monotony effects was also reported by Haider (1956).

Stimulus variety is thought to be important in many theories, such as the one from Hill and Perkins (1985). Scerbo and Sawin (1994; quoted in Scerbo, 2001) asked participants to monitor a vigilance task and a kaleidoscope task differing in stimulus variety. Their expectation was that participants should be willing to terminate the vigilance task sooner, if the lack of stimulus variety led to boredom. They found that participants spent a comparable amount of time on either task if it was their first activity, but 60 % less time on the vigilance task if it was followed by the kaleidoscope task. Subjects felt significantly more stressed, bored, less concentrated, and wished the task to end sooner after performing the vigilance task. But as the authors criticized, results could be interpreted as a consequence of perceived constraints from the experimenter’s side.

Only a few of the studies on boredom included the experience of monotony. Mostly, concepts were operationalized with reports of boredom, fatigue, strain, sleepiness, attention, and irritation as well as with performance and physiological indicators. An often cited study of Frankenhaeuser (1971b) investigated the effect of understimulation (US) in a vigilance task compared to overstimulation (OS) in a sensumotoric task testing simultaneous capacity. Twenty-eight students of a school for policemen participated. In the vigilance task (detection of 16 signals with changing intensity in six consecutive 30-min periods), performance showed a rapid decline over time but improved after ten minutes of rest, whereas in the sensumotoric task (responding to 540 stimuli of different colors and auditory tones), after an initial learning effect followed by a plateau, a gradual decline occurred with no improvement effect after a rest pause. Adrenaline showed a slight increase during US and a pronounced increase during OS. Noradrenaline decreased in the control and US condition, but increased in OS, with a slight decrease after the second run, and comparable to the course of HR. Unpleasantness, boredom and irritation increased over time, with boredom and irritation rated higher in the condition of US. Also, concentration was rated lower in US and decreased over time. High and low arousal groups were formed based on HR and adrenaline. The group with higher adrenaline and heart rate performed better in US while performance was better in subjects with low HR in OS. Unfortunately, it is not known if there was also any relation to the subjective perception of monotony, since both tasks can be considered as monotonous due to their repetitive nature. In addition, the operationalization of understimulation and overstimulation can be criticized. The increased activation in overstimulation contradicts the expectation of decreasing activation under continuous repetitive task demands, but may also be a consequence of constantly required physical activity.

One of the few experiments with the scope of investigating monotony in ATC was undertaken by Thackray et al. (1975). Forty-five participants executed a simulated ATC task for an hour at three different times of the day. They were required to quickly respond to changes in alphanumeric symbols, one important component in the ATCOs task. In the second half of the run, and independent of time of day, heart rate, systolic and diastolic blood pressure, skin conductance level and body temperature were significantly lower while body movements increased. No difference was found in performance and HRV. On a subjective level boredom, irritation, fatigue and strain increased while attentiveness decreased. Two extreme groups with eight participants each of high vs. low boredom/monotony were formed. From the first to the second half of the task, HRV decreased for the low group and increased for the high group, which also had longer response times, a greater decrease in attentiveness and increased strain. Shortest latencies decreased for both groups over time while longest latencies increased for the high group but decreased for the low group.

A psychophysiological approach was also selected by Braby, Harris and Mur (1993) to assess the relationship between subjective components and physiological indicators in underloaded situations. Sixteen male students monitored the flight instruments during two 30-minute-periods. Subsequently, the subjects were split into two arousal groups according to increasing or decreasing HR. There was no difference in HRV, whereas cognitive arousal decreased significantly for the low arousal group. Workload ratings ranged from 1 (low) to 4 (high) in decreased arousal, but from 1 to 3 in increased arousal. Subjectively, more underload was experienced in the decreased arousal group while no significant change or particular pattern for increased arousal group was found.

For a long period, researchers tried to clarify the impact of arousal and activation without finding a clear answer. Davies et al. (1983) summarized that a decline in the level of activation and efficiency has frequently been observed in vigilance situations, pursuit motor performance, simple addition, tracking, and serial reaction time and thus explained the contradictory results. But, apparently not all relevant indicators were assessed in experiments to justify such conclusions. A typical example is the experiment of London, Schubert, and Washburn (1972) that assessed the skin potential level to understand the development of arousal in boredom and added HR in a second study. In both experiments, increased activation was found in a boring monitoring task or letter writing task compared to an interesting story writing task. However, as the tasks were generally rated as tedious, satiation effects might also explain the effect.

Increasing arousal might be a consequence of short-cycle jobs that are often termed hectic and constitute a high workload (Melamed, Ben-Avi, Luz and Green, 1995a). The factor *time pressure* was also discussed by Thackray (1981b). Another consideration is the impact of constraint, but no relation to satiation can be drawn since appropriate indicators were not collected. Hill and Perkins (1985) explained the controversial findings in physiological parameters in combination with boredom with its dependence on mental load during the task. Under high workload but not under low, Perkins found a reduction in HRV in one of his own experiments. In boredom, a person can reject the requirements of a task and stop processing relevant information, or do that intermittently. Consequently, performance declines. But when the person rejects the task, mental load is reduced and HRV increases.

Compared to the amount of lab studies, less experimental studies investigated monotony in the field. A reason may be the difficulty of investigating this issue in a systematic way. In general, studies focused on monitoring tasks such as those executed by control room operators and on assembly line workers. Johansson and Sanden (1989) compared actively planning with passively monitoring control room operators. In a subgroup of 31 operators, monotony was negatively correlated with adrenaline excretion, heart rate, systolic blood pressure, and the perception of work as fun and stimulating, but positively related to the excretion of cortisol and perceived uneasiness. Johansson, Cavalini, and Petterson (1996) also reported a study where adrenaline increased after both passive and active monitoring.

They investigated 24 participants in a 2.5 hr supervisory monitoring task. Performance was more stable after active monitoring and an unpredictable perceptual conflict task was associated with less effort compared to passive monitoring. Finally, higher monotony and less fun and stimulation were experienced in passive monitoring. The increase of cortisol in relation to increasing traffic load and perceived workload has also been observed in 158 ATCOs (Zeier et al., 1996). However, it is noted that from these studies the transfer of the stress concept to explain underload in ATC may be critical because of time pressure reported by the active monitoring group and the lack of freedom to organize their work. While the former is rather obvious in high traffic load situations, the latter might be less relevant in ATC, where the ATCO is responsible for the working methods he or she uses.

A study of a partial sample of 27 out of 232 Austrian ATCOs revealed a significant increase in mental fatigue and satiation towards the end of a shift (Hoffmann & Lehnert, 1992). With increasing complications during work, fatigue is consolidated, while monotony decreases. In the interpretation of these questionnaire results it needs to be considered, that the assessed states are known to partly covary.

Summarizing research results, effects of uneventful and repetitive tasks have been found in several indicators on multiple levels. That the multilevel assessment has been deployed in few fully controlled experimental conditions explains the partly contradicting results. As extensively described in the last section, on the physiological level, monotonous conditions resulted in changes in heart rate, heart rate variability, blood pressure, muscle activity, oxygen consumption, and hormone excretion. Also, indicators for electrooculographic, electroencephalographic and electrodermal activity reflected changes. Subjectively, repetitive and uneventful conditions were often associated with sleepiness and boredom, unpleasantness, irritation, reduced attention and concentration. Again, motivational and cognitive components need to be distinguished. On the behavioral basis, monotonous working conditions resulted in increased errors and reaction times and a higher variation in performance indicators. Haider (1956) discussed a general loss of capability to respond to all peripheral information sources, including the ones relevant to machine operations. The impact on injuries and accidents was mainly analyzed in assembly line workers. Branton (1970) reported that a decrement in automatized motor skill proficiency leads to injuries. Analyzing the accident frequency as a function of working time, it increased progressively during each working period and reduced after every pause.

Little is known about the long-term consequences of monotony, as interactions can occur through complex paths. Health problems might be mediated by the influences of shift work and other organizational factors related to repetitive tasks and underload. While controlling for these factors, Melamed et al. (1995a) investigated the impact of repetitive work and work underload typical for monitoring tasks on coronary heart disease risk indicators (SBP, DPB, cholesterol, plasma glucose) in 2776 male and female blue-collar workers. They found that repetitive work, and to a lesser extent, underload, were associated with risk indicators. This was especially pronounced for women in short-cycle repetitive work, which is usually described as hectic. The mentioned studies which recorded cortisol and catecholamines seem to point in the same direction, even though no effect on BP was reported. Generally, underemployed workers also report lower levels of health and well being as a consequence of their skills and status (Frieland & Price, 2003).

2.4.2. Contextual and Organizational Factors

A definition of Dey, Abowd, and Salber (2001) describes context as any information that can be used to characterize the situation of an entity, which is a person, place or object relevant to the interaction between the individual and the task. Moreover, a task is generally embedded in a work organization within a certain invariable environment and a variable situation.

Concerning the relevance of environmental factors, high room temperature (Barmack, 1939b; Bartenwerfer, 1957), soft light (Bartenwerfer, 1957), and uniform noise (Bartenwerfer, 1957, Voelk, 1988) or a quiet environment (McBain, 1961) had a favorable impact on monotony. In contrast, Mavjee and Horne (1994) found no effect of room temperature in boring conditions. This result can be explained by the strong emotional involvement in boredom, which is not comparable to monotony. The role of a monotonous road environment compared to a more varied one was investigated by Thiffault and Bergeron (2003b) in simulated car driving. Monotonous visual stimulation resulted in more frequent large steering wheel movements which indicates fluctuating performance.

Further factors concern the organization of the task such as the work-rest break design and the possibility of changing the task, which is contained in the concepts of job control. Time effects comprise time-on-task, time-at-work, time-since-awake and time-of-day (Gaillard, 2003) and are of a high relevance in laboratory studies. Reduced signal detection has been found to occur over the first 20-35 minutes (summarized in Edkins & Pollock, 1997). Monk and Leng (1982) confirmed the time of day effects in simple repetitive tasks involving motor activity. Bartenwerfer (1957) found that monotony was more likely in the afternoon than in the morning. The effect of sleep loss/deprivation was affected by task duration (Caldwell & Ramspott, 1998) and time at work (Gillberg & Akerstedt, 1998) and resulted in strategy changes and subsidiary tasks (Hockey, Wastell, & Sauer, 1998), increased sleepiness, and reduced alertness. Smulders, Kenemans, Jonkman, and Kok (1997) also reported differential effects depending on age. However, Gillberg and Akerstedt (1998) did not find physiological correlates related to misses in a vigilance task and thus could not support the idea that they occur because of drowsiness or sleepiness. In a study conducted in a simulated thermal power plant, performance was not negatively affected when operators worked a day- and a nightshift, despite increased sleepiness during the night shift (Gillberg, Kecklund, Goeransson, & Akerstedt, 2001). While similar effects have been found on the subjective level in ATCOs (Luna, French, & Mitcha, 1997), impaired performance was reported also during the night shift (Luna, 1997).

In the field the time factor is usually addressed in the organization of the work shift. Work scheduling and rostering are important factors in jobs requiring 24-hour manning and are marked by many uneventful periods with less work to do. At the same time, some centers are dealing with a high amount of enroute traffic from long-distance flights at certain periods, which might represent rather repetitive traffic patterns. Lille and Cheliout (1982) investigated the differences between diurnal and nocturnal waking states in 22 French ATCOs and found increased HR during night shift. In EEG, the slow delta waves significantly increased at the end of the afternoon and during the second part of the night. This contradicts the results of Costa (1993) who reported decreased HR in the night, but is in line with the results of Melton, McKenzie, Polis, Funkhouser, and Iampietro (1971). The effects of shift length, shift rotation and schedule systems were addressed in various studies (Boquet, Cruz, Nesthus et al., 2002; Della Rocco & Cruz, 1995, 1996**;** Hennig, Kieferdorf, Moritz, Huwe, & Netter**,** 1998**;** Macdonald & Bendak, 2000**;** Melton & Bartanowicz, 1986; Ognianova**,** et al., 1998; Roach, 2003) and revealed a relationship between alertness, performance, and wellbeing.

Constraints and control or degrees of freedom can be understood as opposing concepts. To further investigate the role of constraints in boredom, Prinzel, Sawin, and Scerbo (1995, quoted in Scerbo, 2001, p. 272) repeated the experiments of Scerbo and Sawin (1994; *cf.* 2.4.1), but asked participants to complete each task. In this experiment, the vigil task was compared to the kaleidoscope task and experiences were rated significantly higher in stress, boredom, irritation, sleepiness, difficulty in concentration, and the desire for the task to end sooner. Ratings decreased in alertness and relaxation. Also, the ratings were higher in the constrained conditions, where participants could not terminate the task at will, than in the unconstrained conditions. So the time-constraints imposed on the participants contributed to the feelings of boredom and stress. This aspect is also similar to the continuous allocation aspect in Bartenwerfer's theory and closely related to the degrees of freedom one has in a task.

As a final remark, due to the demonstrated influence of situational variables, it is evident to further distinguish between contextual monotony, which can be caused by the environment, and task-related monotony as a result of the task demands.

2.4.3. Individual Factors

The possible influence of individual factors has been explored in the context of studies on monotony. In early studies, researchers had a major interest in extraversion/introversion, even though contradictive results did not support a clear relationship with performance in monotonous tasks (e.g., Bartenwerfer, 1957; Koelega, 1992). The investigation of 63 female workers at press-operating jobs revealed significant correlations of reported boredom and fluctuations of boredom with age and neuroticism but none with extraversion, intelligence, length of service and work variety (Hill, 1975). Brocke, Tasche and Beauducel (1997) observed that extraverts and introverts indeed react with different levels of compensatory effort when confronted with low or high degrees of stimulation. In driving behavior Thiffault and Bergeron (2003a) found that sensation seeking predicted drivers' performance. An interaction between sensation seeking and extraversion implied a tendency that those who have high values in both scales perform worse.

The tendency to be bored has been expressed in the concept of boredom proneness. Boredom prone individuals show higher impulsitivity, lower attributional complexity, lower need for cognition, are less sociable, less assertive and more alienated (Harris, 2000). Boredom prone individuals have higher scores on negative and lower scores on positive affect measures (Vodanovich, Verner, & Gilbride, 1991). Boredom proneness predicted cognitive failures and was associated with daytime sleepiness scores (Wallace, Vodanovich, & Restino, 1993); also crash-related conditions and risks in driving were predicted (Dahlen, Martin, Ragan, & Kuhlman, 2004). Watt and Blanchard (1994) found in 214 undergraduates that individuals high in the need for cognition, defined as the likelihood that individuals engage in effortful cognitive experiences, are less prone to experience boredom. In 170 college students boredom proneness was positively correlated with mood monitoring, i.e. the tendency to direct the attention to ones mood, and negatively with mood labeling, i.e. the ability to direct ones mood, and the experience of flow (Harris, 2000). It is noted that these concepts are relevant, as effort is necessary to direct ones attention to a task perceived as boring. High levels of job boredom and boredom proneness significantly relate to lower job satisfaction scores and organizational tenure and greater absenteeism (Kass, Vodanovich and Callender, 2001). O'Hanlon (p. 63) reports several studies, where subjective boredom, job dissatisfaction or both diminished as a function of age for both male and female assembly line workers. Boredom prone fire-fighters perceived their jobs as underutilizing their skills and abilities, having greater organizational constraints and a lack of variety (Watt, 2002). Vodanovich, Weddle, and Piotrowski (1999) reported that individuals high in boredom proneness possessed greater external work value scores (attitude towards earnings, social status, upward striving) while the ones low in boredom proneness reflected higher internal work values (pride in work, job evolvement, activity preference). Kogi (1972) reported a relationship between staff indifference, as an attitude to their job, and greater vigilance decrement in two electricity power firms. Van der Flier and Schoonman (1988) even found that previous accidents and less job satisfaction were predictive for future train accidents.

A similar concept is susceptibility to industrial monotony. Smith (1955) questioned 72 woman workers in a knitwear mill who were engaged in light repetitive work. Monotony-susceptible workers were described as young, having a tendency to day-dream, more restless outside the plant and more intelligent. There was however no relation to introversion-extraversion. In addition, large individual differences in susceptibility to drowsiness were also observed, which may relate to monotony susceptibility (Verwey & Zaidel, 2000).

Little is known yet about action control style and monotony, but it might have an influence on the development of strain, since state oriented persons focus attention differently when they are confronted with a demand (Bossong, 1999). The construct of action control style has been introduced to build the gap between choice and action (Kuhl, 1994a). Action orientation is the ability to facilitate the enactment of context-adequate intentions. If debilitating effects on individuals volitional abilities to plan, initiate and complete intended activities are found, state orientation is dominating. Consequently, the intrusive and perseverating thoughts while focusing on one's state can impair volitional functions. The impact of intrusiveness on vigilance performance was demonstrated in a study of Schapkin and Gusev (2003, p. 148). The withdrawal from an uninteresting or demanding task towards personal concern is especially a risk in monotonous settings. The authors also discussed the possibility that the processing of the negative affect, as predominant in uniform work settings, transforms into intrusive thoughts. Intrusiveness might also be negatively related to expertise, as it might depend on the available strategies ATCOs dispose of to counteract monotony. A small indication for this assumption is given by Idogawa (1991) who found that experienced drivers were less drowsy as indicated in alpha of EEG compared with untrained participants. This was interpreted as indicator for a well-developed self-control; however, the number of participants was very low (n=5).

In summary, it turns out that the individual differences in confronting a task might be an important factor to prevent monotony. Other concepts like morningness-eveningness have not been found in the literature. Even though it might be important as relationships with the initial state might be assumed, its relationship with monotony has not been measured yet. Despite inconsistent results in traditional factors, the recent interest in additional individual differences underlines the consideration of such factors in selection research.

2.5. STRATEGIES TO AVOID OR MITIGATE MONOTONY

This section investigates different approaches on how to avoid or mitigate monotony. Each of them has its advantages and disadvantages, which have to be considered carefully for an eventual transfer to the area of ATC. Approaches distinguish countermeasures according to their impact on task conditions, working environments or individuals. Furthermore, it is important to differentiate the most likely occurring critical controller states to identify the most appropriate countermeasure. Even though fatigue and satiation might be closely related to monotony as far as it concerns the described effects, the most appropriate countermeasures for each state are different. Herein, in agreement with the socio-technical approach (*cf*. Badham, Clegg, & Wall, 2000), countermeasures are allocated to act predominantly at the task, individual or organizational level. A distinction between prevention (action before the task starts), intervention (the problem already occurs and can be alleviated) and retro-active strategies (the problem has occurred but improvements can be made to avoid it in the future) is considered (Gaillard, 2003), also known as primary, secondary and tertiary prevention.

2.5.1. Strategies Affecting Task Design

Strategies assigned to this category mainly have the goal to optimize a task according to the human limitations. Job enrichment, enlargement and variation are seen as efficient countermeasures against monotony (Ulich, 2001), regarding the stimuli to be processed as well as the required actions for task completion. The effect of increasing variety seems to be dependent on the number of times that different tasks were alternated (Wyatt, 1929). Walker and Guest (1952; quoted in Davies, 1983) interviewed assembly line workers and found that with a greater number of executed operations the interest in work was rated higher. Bartenwerfer (1957) reported the positive effects of increasing work speed.

Nonetheless, job enrichment does not completely resolve the issue (Davies et al., 1983). Karsten (1928) observed performance deterioration in a repetitive task, but recovery occurred rapidly with a change in the task set. Sudden shifts between high and low workload situations may have a bigger influence on the work than high or low workload itself (Smolensky, 1990). But unexpectedly, Krulewitz, Warm, and Wohl (1975) observed a pronounced performance decrement in monitoring, if the slow event rate was shifted to a fast one. In contrast, performance improved if a fast event rate shifted to a low one. This was explained with a change in the willingness to attend to the task after the demands increased unexpectedly. On the other hand, Gereb (1968) found that monotony was having less impact on performance degradation after an activating task.

One improvement can be reached through providing feedback. Hitchcock et al. (1999) found positive effects of cueing and knowledge of results on workload and boredom in sustained attention. On the other hand, interruptions may have a negative effect on the experience of boredom (Fisher, 1998). In two studies with a total of 352 students exposed to different tasks and interruptive conditions, external interruptions reduced boredom in simple tasks requiring little attention, but had no effect on simple tasks requiring attention or complex tasks. Internal interruptions in the form of non-task-related thoughts about personal concern were related to higher boredom and less satisfaction at work. One solution is to provide people with control, as Scerbo, Greenwald, and Sawin (1993) investigated in a vigilance task. Participants who could control the event rate did not show the decline in performance, even though the high workload ratings did not differ. Hockey (2003, p. 12) reported that people perceiving controllability in the work environment appear to have a buffer against strain and performance decrement. Less negative effects of work on health and well-being occur, as they can more effectively manage their state.

The impact of instruction was observed in several experiments as it may influence how individuals conceptualize the task. Already Barmack (1939d) demonstrated that the expectation of how long someone needs to work had an impact on output. In adding numbers participants performed faster in the first hour if they were scheduled to work one hour than the ones working up to four hours. At the same time, the subjective feelings evolved in a similar pattern in all conditions independent of the time-on-task, leading to the assumption that motivation and energies are shared differently. In an experiment of Sawin and Scerbo (1995), one group received a traditional detection-instruction, while another group was requested to evaluate a special display. The control group was asked to relax in front of blue light. Receiving a detection-emphasis instruction resulted in higher workload than receiving a relaxation-instruction. In 30 minutes vigil, there was no difference in the performance decline, but the ones who received the traditional instruction reported higher levels of workload. Also, participants low in boredom proneness outperformed participants with high values. The instruction manipulation did not affect the overall boredom rating, but did influence the stress rating. Thus, this was supporting the importance of cognitive appraisal of a task as it is also included in Hacker’s (2003) action regulation theory, the cognitive evaluation in the models of Hill and Perkins (1985) and Matthews (2001), or the importance of giving a personal meaning to a task (Oboznov, Yegorov, & Kostritsa, 1991). Additionally, as Hukki and Norros (1990) state, the cognitive interpretation of the situation in the course of the action is also essential, not only the action itself.

The consideration of task involvement is another approach related to automation issues (Riley & Parasuraman, 1997) and was addressed in studies comparing active and passive monitoring (e.g., Perdson, 2001; Metzger & Parasuraman, 2001). Willems and Truitt (1999) investigated the effect of active controlling versus monitoring in 16 ATCOs and found that situation awareness was lower in the monitoring condition despite reduced workload. This was even more pronounced if traffic load increased. A more recent trend is to develop devices that help to recognize critical states. Verwey and Zaidel (1999) proposed an alertness maintenance device to prevent drowsiness in driving accidents. However, as a secondary prevention measure it only acts after alternatives have failed.

But, as research on adaptive automation has demonstrated, until today research on predictive indicators has not been as successful as expected for a long time (*cf.* Wilson & Russell, 2003b). As a final remark, in spite of the fact that a lot of research has been conducted in laboratory environments, with a foresighted design of a task a lot of problems can be anticipated from the operational environment.

2.5.2. Strategies Affecting the Working Context

Strategies within this category comprise work scheduling factors as related to the length of work periods and shifts, rostering, management of frequency and length of rest-breaks, and the organization of the environment. Even though a consideration of these aspects primarily addresses fatigue, a balanced organization might also affect monotony.

The most important feature of rest breaks might be seen in counteracting fatigue and avoiding the accumulation of strain through determination of appropriate work periods. In ATC related experiments as well as in ATC centers working periods do generally last up to 2 hours. Depending on local arrangements, centers introduced a minimum of 30 minute rest breaks. However, work period duration has hardly ever been based on empirical results obtained in field studies and regulations vary between countries. Also, the recommended time at a position might depend on the individual state as well as on the traffic characteristics. Rest break organization is especially a critical subject in ATC, where work is organized in teams. In addition, the sector handover to replace an ATCO has to follow a certain procedure, and the time to build up a mental picture may take around five minutes, hence influences the frequency of rest breaks. This does however not impede ATCOs from taking micro-breaks at the position. In the long tradition of trade unions, a rest break has always been seen as a passive state where the employer has no right to interfere. Conversely, positive effects of controlled breaks have already been demonstrated in several studies (e.g., Neri et al., 2002). Optimal recovery is not only a function of the time spent on a break but on the activity in this time as well. It has been confirmed in different studies that rest breaks do not only support the state to return to the baseline level after exposure to load, but rest breaks also have a function of balancing the prior activity to ensure homeostasis (Luczak, 1998). An appropriate homeostatic energetical state helps to maintain motivation, and interferences between cognition, motivation and energetics are well known.

Bevan, Avant and Lankford (1967) investigated the influence of interpolated active and inactive rest breaks on the vigilance decrement. They showed that introducing five minutes of rest breaks after 30 minutes of work led to a higher performance in a 2-hours-vigilance task than two hours without a break. The break activities were physical exercise, mental exercise and a deprivation condition without stimulus. Kozena, Frantik, and Dvorak (1996) investigated relaxation, physical exercise, intensive light exposure and adding a mental task to enhance the psychophysiological state. These conditions were introduced in the course of three 90-minute-sessions to execute an acoustical discriminatory task and a continuous visual task (n=118). A favorable effect of self-instructed relaxing was found in feelings of sleepiness; the conditions of the additional mental task and light exposure resulted in improved performance and reduced sleepiness. No effect of physical activity (knee bends) was shown, which might be explained with their intensity. This assumption was further investigated (n=21) by Oweis and Spinks (2001) who reported a negative effect of intense physical activity when compared to light intensity, as it decreased subjective energetical arousal and increased tense arousal; however, no effect on reaction time performance was found. Participants who were exposed to sleep deprivation and atropine showed reduced performance in an aircraft identification and vigilance task (McLean, Smith, Hill, & Rubenstein, 1993), without an impact of moderate exercises.

Nonetheless, an effect of moderate exercise has been demonstrated in an everyday-life environment, when Thayer (1987) compared sugar snacking with a fast 10-minute-walk. Snacking increased subjective energetical arousal after 20 minutes. Subjects were more tired and reported higher tensions one and two hours after eating the snack. Alternatively, the fast walk was associated with increased energy and decreased tension during two hours. It was concluded that the 10-minute-walk would be a preferred alternative to increase the energy-level in everyday life. Another explanation for the positive perception of exercises is their impact on mood (Lane & Lovejoy, 2001). Thayer, Newman, and McClai (1994) report that exercise is ranked near the top among the behaviors the respondents use to self-manage their moods. Acute exercise transiently reduced self-rated anxiety, tension, or negative affect (Landers & Petruzzello, 1994) and increased positive affect (McAuley, 1994). However, Crabbe and Disman (2004) criticized that in many studies which used cortical indicators to reflect decreased cortical activation indicating fatigue, relaxation or decreased anxiety, experimental artifacts may have occurred. For this reason they conducted a meta-analysis of 18 studies (n= 282) and found that alpha activity increased most after several short exercises, immediately after an exercise and when exercise was compared to changing indicators after resting conditions.

The introduction of incentives is another way of creating interest in a task. In a study with 36 subjects it was investigated if making a task more motivating or counteracting the hypnotic trend during boring work interfered with the development of a bored attitude to the task (Barmack, 1937, quoted in Barmack, 1939). In executing the task of adding pairs of six-digit numbers the use of 10 milligrams of benzedrine sulphate retarded the development of an unfavorable attitude towards the task. However, the incentive condition as manipulated by Barmack (1939b) is questionable, as he did not introduce a control group, even though monotony was also found to be a mediator in the effect of pay systems on emotional distress (Shirom, Westman, & Melamed, 1999). Gereb (1978) found that monotony was alleviated by competition-like tasks that were performed in teams. This had even a stronger impact than a change of activity. Bartenwerfer (1985) also suggests enabling social activity and communication. However, few experiments were considering the social factor. Johansson et al. (1996) did not find a favorable effect of social contact on performance but noted a positive attitude towards the presence of a second person. Concerning the impact of music, Gereb (1968) reported that despite its positive perception by participants in a monotonous task and its recovering effect, it affects performance negatively because of its distractibility.

2.5.3. Strategies Affecting Operators

Opposed to designing an ideal task for the operator (*cf.* 2.5.1.), this section discusses ways to select the ideal person for executing a defined task. The selection and assessment of individual characteristics and capabilities is of major importance, as it occurs at an early stage before a person is in the situation to actually execute a task. Especially when the task cannot be adapted any more and options to control are rare, the selection of the appropriate operator is essential.

This leads to the question of which individual indicators predict critical states. As Touzé (2005) pointed out, personality appeared to be a valid predictor of performance when valid instruments were used and the links among constructs were systematized. The Five-factor model, frequently referred to as the Big Five (e.g., Digman, 1990), represents a common framework with conscientiousness and emotional stability having a strong impact. Extraversion and conscientiousness were also tendencially correlated with the performance in a short vigilance task executed by 96 participants, while neuroticism was rather related to perceived frustration (Rose, Murphy, Byard, & Nikzad, 2002). As a consequence, Bartenwerfer (1961) recommends the exclusion of persons with high extroversion.

This approach does not consider that after selecting a person, it is still possible to assign a controller to selected positions depending on the expected traffic characteristics. This balances the variation in individual preconditions on a day-to-day basis; however, no empirical support has been offered yet.

Training and the formulation of recommendations are relevant issues in how to cope with situations to change behavior. An example is the training of how to maintain situation awareness as it was proposed by Redding (1992). Similar efforts in the field of ATC have also been undertaken by the FAA in tackling the issue on how to counteract the influence of fatigue. However, recommendations concerning monotony in ATC might be different, but have not yet been formulated and empirically supported. This also goes along with an early recognition of critical states in oneself and others within the working team, and is followed by how to avoid such states and how to cope with them once they emerge.

2.6. SUMMARY

An extended analysis of the literature and theories that might be relevant for the current research was undertaken to develop an approach that might be relevant for the investigation of monotony in ATC.

Starting from monotony as a phenomenon that occurs in uniform work environments, it became clear that monotony has formerly been used to designate the task characteristics as well as the individual reactions to these characteristics. This favored a clear distinction between the description of task characteristics in terms of uneventfulness and repetitiveness as well as an exclusive use of monotony for the individual state. But it is also obvious that using related terms loosely does not help to gain more knowledge of the subject. The related concepts of boredom, underload and low vigilance need to be clearly distinguished and kept apart from the critical states fatigue, satiation or stress, which might occur as a further consequence of task execution.

One explanation for the concepts getting mixed up is that they are often similar in appearance. For example, fatigue is compared with empty energetical resources, but, because of this focus, it is not necessarily a consequence of monotonous tasks despite feelings of tiredness. Another basic critical state to be distinguished is satiation. This concept emphasizes emotional involvement, while the above mentioned terms are rather concomitants of these states or generalizing entities. Stress is a phenomenon often associated when working conditions have reduced options for control, but it has however a totally different nature. To well describe the discussed concepts, how to conduct the assessment becomes a very important question. As the state conception supports a complex development within these multiple indicators, the assessment needs to consider this variety. It is the combination of physiological, subjective and behavioral indicators which reflects the state of an air traffic controller.

Moreover, it needs to be considered that, when dealing with monotony, we are dealing with a problem that occurs within a working environment and thus necessitates the consideration of organizational aspects. Thus, previously developed models to explain the effects of task execution on performance and individual states are often too restricted. The issues of recovery and resources are factors that contribute to good performance and hence need to be included. On this basis, resulting strategies to counteract monotony can be distinguished according to a prior focus on the task, the work environment or the individual.

Unfortunately, even less is known about ATC, where only very few researchers have specifically tackled the problem of monotony. One of the pioneers in researching monotony in ATC used monotony synonymously with boredom. Nonetheless, an interesting outcome of the application of a multivariate approach revealed higher physiological deactivation, slower reactions and increased fatigue under high boredom and monotony.

A questionnaire study in the operational field resulted in increased monotony towards the end of the shift, which was however reduced depending on the amount of complications controllers dealt with. More recently, the investigation of consequences of passive monitoring might have brought up the issue of monotony again, but remained focused on workload and situation awareness. Finally, a consideration of the research in related fields supported the definition of the research approach.

3. DEVELOPMENT OF A RESEARCH FRAMEWORK FOR MONOTONY IN ATC

The following chapter presents the framework for the current studies. After a detailed description of the research problem based on the literature background, the motivation for the research is explained. Objectives are deducted and translated into research questions and an expected output. Finally, the necessary research activities to reach this output are described.

3.1. PROBLEM STATEMENT AND MOTIVATION

As was pointed out in the introduction, the issue of monotony has not been systematically investigated in ATC. A variety of problems[4] and limitations characterize this situation and different reasons are hypothesized to explain this undervaluation.

First of all, monotony is an unclear concept and various subjects implicitly relate to it. Because of a lack in definitions that contributed to a common understanding, many researchers applied the term according to their actual needs. The basic problem is that monotony has been used in two meanings, which is the term to characterize objective work conditions and the notion to describe an individual state. Even though researchers have known this problem, and it is similarly contained in efforts to define the concept of stress, in monotony research this has not been clearly addressed. Stress researchers solved this issue with the recommendation to clearly mention the background one has in stress research. In the current work, monotony is applied to describe an individual operator state to avoid the confusion between monotony addressing the task factors and monotony describing an operator state. The descriptors *uneventful* or *repetitive* are preferred to denote monotonous task factors.

Secondly, despite tendencies to create more uniform working conditions in response to increasing traffic demands, traditional approaches on stress have dominated the applied research, and issues of workload and stress management were advanced. One example consists in the introduction of critical incident stress management in different ANSPs (Baumgartner, 2004; Vogt, Leonhardt, Koper, & Pennig, 2004). Ongoing synchronization tendencies, for example operating traffic flows of homogeneous aircraft, suggest a consideration of effects of uniformity.

Finally, the issue of automation in system development has not been directly connected with monotony. Although active and passive monitoring were compared in several studies (e.g., Willems & Truitt, 1999), the assessment was concentrated on performance and related workload. Main approaches to counteract negative side-effects of automation relate to the concept of adaptive automation or the evaluation of tools such as the conflict resolution advisories in ATC (Schaefer, Flynn, & Skraaning, 2003). The strength of the latter is that after proposing conflict resolutions, the final authority is left to the controller. Despite these developments, underload remains a topic in discussion. For example, in the proposed concepts the range of potential activities required to execute the overall task remains reduced. General interest remained scarce within the field of ATC, even if several authors stressed the importance of monotony or boredom. Additional reasons for a limited transfer of previous research on monotony to the field of ATC include:

- In laboratory environments tasks such as simulated driving were performed.
- The nature of assembly line work is different with regard to the responsibility and safety-criticality, the required skills and training and the existence of a normative/optimum task sequence.
- Monitoring or maintaining vigilance is just one of the performance components in ATC, even though it was frequently investigated in numerous studies.

[4] The term *problem* is used in the scientific sense of a question to be considered, solved or answered.

In the next subchapters, the problems related to the understanding of monotony in ATC are further elaborated. Different perspectives analyze task factors evoking monotony, the description of states of monotony or potential countermeasures and mitigation strategies.

3.1.1. Problem context 1: Understanding Factors that Evoke Monotony in ATC

Elements causing monotony have not been well described in ATC. Relevant characteristics and issues were mainly analyzed in related tasks and fields, where a number of factors have been mentioned to evoke the earlier introduced state of monotony. Basically, repetitive and uneventful work conditions are considered in relation to Johansson's (1989) proposal to distinguish uneventful and repetitive monotony. The disadvantage of this proposed framework is that it focused on the description of two different work environments as characteristic for uneventful monotony (supervisory control) and repetitive monotony (assembly line workers), making a clear reference to monotony as an individual consequence and describing differences in task characteristics such as complexity, predictability, work environment, payment, and amount of control. This approach offers a good description of relevant aspects. However, in this thesis the terms uneventfulness and repetitiveness will be reserved to denote task characteristics. Also, it has not been considered in Johansson's framework that uneventful and repetitive monotony might occur within one job, as it is presumably the case in ATC. Thus, the combined impact of the earlier mentioned factors such as complexity, predictability, work environment, and amount of control might play a significant role in the work of air traffic controllers.

Obviously, uneventful monotony can occur in situations of low traffic that require few actions to deal with aircraft. Such a situation varies for regions and centers, but often occurs during night shifts. This aspect has been addressed within vigilance research (e.g., Schroeder et al., 1994) with the argument that such monitoring situations mainly demand sustained attention. Nonetheless, results of vigilance research are not directly transferable to counteract monotony, as was discussed in 2.1.5. Although controllers have to handle little traffic, they need to remain alert to deal with eventual conflicts. Even in situations of low traffic, a certain task complexity is existent, thus the action cycle includes a variety of steps to complete the task goals.

Less apparent is the potential of repeated execution of task elements to lead to monotony, as it was present in the study of Thackray et al. (1975). Repetitive monotony can result in medium or high traffic conditions, if task characteristics do not display a certain variation or if difficulty remains below a challenging threshold. This often happens when the traffic is dominated by so-called "hello-good-bye traffic", e.g., traffic that does not require action to prevent conflicts. Thus, the nature of the traffic has the potential to cause monotony in many centers, as repetitiveness can be found in various forms. Task characteristics include runway allocation affecting approach and departure routes, certain sector forms may reduce the planning span for each aircraft, repetitive flight plans lead to routine traffic, or parallel airways with few crossing points reduce the complexity in projection of the traffic. Monotony can also depend on the distribution of actions, for example the amount of controlling or monitoring. Indicators which are independent on traffic requirements consider the standardization of procedures and homogeneous working methods.

Another argument that supports the distinction between uneventful and repetitive conditions is that the nature of potential errors may be different depending on the condition that evokes monotony. While in uneventful monotony errors may be a consequence of suboptimal activation and consequently slow re-adaptation when action is required, in repetitive monotony errors may relate to omissions in the update of the action cycle. This reminds on Luchins' (1942) mechanization of thought, even though it has not been related to the concept of monotony.

However, this does not necessarily mean that a difference in the pattern of the appearing psychophysiological state of monotony is present. Also, it is not sufficient, as correctly considered by Johansson, to only focus on task factors. Additional factors stem from the work environment and organization, and individual differences may play a role. Depending on someone's state at the beginning of the shift, he or she may react differently to situations potentially provoking monotony. Moreover, a different length of working schedules or the time worked at the positions influences monotony and affects issues of shift organization.

A different view on the situation and individual preferences to act and deal with task demands might have an impact on the resulting state. Table 2 presents a summary of these factors.

Table 2: Factors expected to be relevant for evoking monotony

CATEGORY	FACTOR
Task characteristics	Traffic repetitiveness
	Traffic density
	Time-on-task
	Work procedures
	Planning span/ aircraft time in sector
	Standardization of controllers responses
Individual factors	Initial psycho-physiological state
	Age
	Training/experience
	Personality traits
Work environment	Temperature, noise
Work organization	Shift system

In summary, monotony as a consequence of task execution in an air traffic control environment can be seen as a result of primary task characteristics. These factors may interact with additional individual and non-individual influences. Individual factors include age, experience or training as well as the psychophysiological state at the beginning of work. Personality traits include dispositions for experiencing boredom or a certain preference in working styles related to personality factors such as extraversion or conscientiousness. Non-individual factors comprise characteristics of the work environment (e.g., room temperature) or work organization (e.g., work schedules, work shift). However, the current problem is that a combination of these factors has not been included in study designs to define their role in evoking monotony. This is especially true for ATC.

3.1.2. Problem context 2: The Description of Monotony and Other Critical States

As pointed out, the term *monotony* is used to denote to an operator state rather than task characteristics causing such a state. Monotony is seen in the tradition of Bartenwerfer (1960) as a slowly developing state of reduced activation caused by activities characterized as repetitive, long-lasting, lowly stimulating or lowly difficult. Mainly it is associated with feelings of sleepiness, fatigue, the task is perceived as uniform or boring. A reduced ability to react and adapt to changes can be connected with impaired and fluctuating performance.

An advantage of this approach to monotony via the description of a task and the individual in the interaction with the task is that it allows differentiating between different states. This is also in accordance with Smith (1955), who emphasized that repetition is defined by the “externally observable” frequency of the occurrence of an event. It cannot be identified as *the* cause of monotony due to the importance of cognitive interpretation as stressed in different models. Individuals have various strategies to deal with situations as well as different cognitive interpretations. For this reason, a description of monotony needs to include the individual's subjective interpretation, as considered by Bartenwerfer (1957), McBain (1970) and others.

Explanations of how monotony develops in ATC can be deducted from models on workload or strain, which were developed in other fields. Different models proposed to explain the individual effect of the work process have their shortcomings as they either ignore energetical or cognitive processes, traits, or changes over long-term periods. They seem to be applicable only to formerly defined characteristics and cannot be generalized. For example, there is no global relationship between high traffic load and stress, and also stress as the result of low traffic may only be true in very limited cases. Moreover, it has not been considered that it depends on the type of high traffic load in relation to the sector complexity and the variety possible in dealing with this traffic. Thus, the previously used concepts might have been incompatible to describe the effects of traffic load. Under this precondition even the assumed continuum between overload-underload according to McGrath (1970) may not be suitable as it does not consider resources to prevent stress as a consequence of overload. The advantage of action regulation theory is that with the explanation of monotony as a lack of subgoals and hierarchically or cyclically incomplete action cycles it is applicable for both uneventful and repetitive situations at work. Also, Zapf's (1995) distinction between complexity and variety depending on the horizontal level of subgoals or vertical action cycle is useful for air traffic control, as it does not necessarily relate monotony to low complexity. This view can be completed by the assumptions of MART and habituation theory, where the organism adapts to the requirements. Other models help to explain contradicting results, such as the effort loop integrated in control theory to explain maintenance of performance, which supports their integration in a common framework to explain monotony in ATC.

The model used as a background for this thesis is based on the stress-strain-model introduced in section 2.2 under special consideration of the assumptions of action regulation theory. Its strengths are the distinction of several critical states as strain consequences. But it is challenged if the assumption of apparently independent critical states can be maintained as well as the implicitly contained stimulus-response assumption. It might also be the case that different critical states are not excluding each other. This focus was set by Leonova (2003) who understood functional states as a composition of different affective, cognitive, and behavioral patterns.

As pointed out in the ISO 10075 for mental workload, seeing monotony as a fatigue-like state clearly needs to be distinguished from mental fatigue or other fatigue-like states such as mental satiation. Although similar in appearance, as tiredness, boredom and performance degradation are related to each other, in reality the causes are different and thus the implications for work design and organization. The most important distinction to note here is that, unlike fatigue, the effects of monotony as well as satiation can be alleviated by changes in the operator's task, whereas recovery from fatigue effects requires physical and mental recuperation. While in fatigue and satiation energy consumption is increased, in monotony energy consumption is decreased (Richter & Hacker, 1998). Depending on additional factors, states of satiation and fatigue alternatively emerge.

Table 3: A distinction of monotony and similar states in ATC on different assessment levels

EFFECTS		MONOTONY	SATIATION	FATIGUE
Physiological level	Cardiovascular activity			
	Heart rate	↓	↑	↑↓
	Heart rate variability	↑	?	↑
	Electrodermal activity	↓	↑	↓
	Cortical activity	↓	?	↓
	Endocrine activity	?	?	↑
Subjective level	Affective response (boredom, irritation)	↑	↑	↓
	Cognitive functions (concentration, attentiveness)	↓	?	↓
	Motivation	↓	↓	↓
	Sleepiness, drowsiness	↑	↑	↑
	Strain	↓	↓	↑
Behavior	Task performance	↓	↓	↓
	Variability in task performance	↑	↑	↑
	Behavioral activities (e.g., gestures)	↑	↑	↑↓
Long-term effects	Work satisfaction	↓	↓	-
	Health impairments	↑	↑	↑
	Absenteeism	↑	↑	-

Furthermore, we can resume from previous studies that mental strain may depend on task characteristics, individual factors, and the organizational context. At the same time the occurrence of positive states such as flow, when someone enjoys the task, cannot be ignored. Thus, the application of this approach would especially be useful for ATC, as in operational settings the occurrence of failures is frequently attributed to fatigue, without any further distinction from fatigue-like states.

The distinction of various critical states in the context of monotony requires their accurate assessment, for which the psychophysiological approach is considered appropriate. A combination of subjective, behavioral and physiological measures in a multivariate design ensures the correct interpretation of the reaction pattern. It becomes evident that the distinction between the states is very difficult as differences may be observable in very few indicators. As a summary of research (compare studies in Chapter 2; e.g., Bartenwerfer, 1957; Barmack, 1939a,b; Johansson et al., 1996) Table Error! Not a valid link. specifies the main indicators which can be used to assess monotony and which have the potential to differentiate the critical states of monotony, satiation or fatigue. Some indicators cannot be clearly assigned to a certain state based on previous research. Also, long-term effects and the impact of motivational factors are difficult to demonstrate. Under consideration of these aspects, the multivariate approach is promising for an application in the domain of ATC.

3.1.3. Problem Context 3: The Development of Countermeasures

As it became clear in the last chapter, potential countermeasures for dealing with monotony and other critical states can act through a variety of paths, but neither of them has been systematically investigated in ATC. Also, strategies that address the task design, the work organization and the individual are distinguished. Moreover, countermeasures for ATC are defined for all implementation levels regarding the temporal perspective that distinguishes initiatives for prevention before an activity is executed, intervention during task execution or retro-active strategies. Which countermeasure might be most effective in which condition depends on the causes of monotony. In situations characterized by uneventful tasks the proposed methods might be different compared to situations with high requirements for uniform and repetitive actions. The strategies might be applied by different concerned parties, for example, the organizational management, the supervisor on shift, the system designers or the ATCOs. Also, some of them may have an additional positive effect on the long-term and are expressed in health, well-being or work satisfaction. However, for all the proposed strategies the specific characteristics of a highly safety-critical work environment need to be kept in mind. A selection of strategies which are principally applicable is contained in Table 4.

Table 4: A collection of potential strategies and countermeasures to mitigate monotony in ATC acting upon different levels

LEVEL	PREVENTION (PRIMARY)	INTERVENTION (SECONDARY)	RETRO-ACTIVE (TERTIARY)
Task	task involvement	instruction	task re-design
	action variety	feedback	
Organization	design of work environment	rest break planning	correction of work environment
Individual	selection	training	apply strategies

Some of the countermeasures are easy to implement through individual contributors, while others require the involvement of management authorities. The consideration of potential side-effects on the primary level as for example in early task design and concept development is seen as the preferred alternative. The complementation with additional strategies is applied once certain situations have been perceived as critical with regard to safety and thus do present only secondary options.

3.2. RESEARCH OBJECTIVES AND RESEARCH QUESTIONS

To tackle the limitations and problems described above, the overall objectives of this research to be dealt within a systematic investigation are:

- To find out how a state of monotony can be operationalized in ATC, thus, which factors do evoke it and how it can be measured.
- To look whether repetitiveness and uneventfulness result in similar states.
- To identify additional contributing factors.
- To determine countermeasures that may be introduced to reduce the potential of negative side-effects.

If factors were known to be relevant for monotony, they can be systematically included in both the development of future ATC concepts and tools as well as in the current and future organization of the operational environment. Increased knowledge on monotony in ATC supports the implementation of the most suitable ways to deal with critical states. In consequence, it contributes to increased air safety through reduced contribution of the factor monotony in critical situations. In the following, the main objectives are presented which address the described problems and research questions are formulated.

3.2.1. Objective 1: The Definition of Factors Evoking Monotony

A systematic investigation of factors that evoke and contribute to the development of a state of monotony in air traffic controllers is needed. Since research results from related fields cannot be directly transferred, a systematic investigation of these factors in the ATC environment is necessary.

A central assumption is the role of uneventful and repetitive work conditions whose combined function has not been investigated in ATC environments despite an expected impact. Different factors may be of interest in such a situation, and therefore a focus on pre-selected characteristics is preferable. Task repetitiveness is seen as a relevant factor, while a second aspect addressed the importance of traffic density in relation to uneventfulness. Even if repetitiveness was a dominating factor, it might act differently under low and high traffic density. The concept of dynamic density (DD; e.g., Laudeman, Shelden, Branstrom & Brasil, 1998) allowed the consideration of task difficulty especially in the simulation-setups without being restricted to the operationally applied *rule of thumb,* the aircraft load. Thus, if these factors were confirmed as significant contributors, their analysis would imply their consideration in concept development. However, as additional factors might come up during the course of the research, the research question is generally stated:

<u>Research Question 1.1:</u> Which factors can be defined in the task of an air traffic controller that allows the operationalization of an individual state of monotony?

In a next step, it is asked for additional factors on an individual or contextual level that mitigate or reinforce the effects of repetitiveness:

<u>Research Question 1.2:</u> Which factors aggravate or reduce the development of a state of monotony?

One such factor might be the shift in the traffic load from low to high or from high to low traffic within one work period. For example, it might be possible that a stress-like reaction pattern emerges, which overlays monotony when a high traffic load condition is followed by a low traffic period. It is however not known how long this effect lasts. On the other hand the increase of traffic density may act as a counteracting factor after low traffic. Further factors concerning individual, trait or state factors were included, as a potential determination as significant contributors would help to understand which additional influences mediate monotony.

3.2.2. Objective 2: The Description of Monotony and Other Critical States

Monotony as an individual state can best be described through the assessment of indicators from multiple levels containing physiological, subjective or performance measures. It is assumed, that a combination of different indicators might be most successful for defining monotony. However, it needs to be determined in a first study which indicators are the most appropriate. Based on this investigation, they can be summarized into an indicator for the state of monotony which is investigated in the following studies.

The importance of appropriately assessing different critical states has been underlined with respect to the introduction of the most efficient countermeasure. Also it is not clear if the states develop differently over time. The following research questions are thus addressed within this objective:

Research Question 2.1: Which physiological, subjective and performance indicators can be defined that allow the description of the state of monotony in air traffic controllers depending on the traffic characteristics repetitiveness and dynamic traffic density?

Research Question 2.1: Which indicators distinguish best between monotony, fatigue, and satiation in different phases during task execution?

3.2.3. Objective 3: The Definition of Countermeasures and Strategies

After describing the situation, the final question addresses how to improve it. A variety of strategies acting on the different factors have been shown to have an effect on monotony, but only few have been systematically investigated. Countermeasures can principally affect task design, the individual, and the work environment (see Table 3). Partly, the factors of interest are selected depending on prior research and with regard to their relevance in ATC. In previous research, a positive effect of physical activity in short rest breaks during the execution of an air traffic control related task was found (Straussberger & Kallus, 2003). Thus, systematic application of break exercises might not only be relevant to decrease fatigue, but also to reduce negative side-effects of monotony, and is considered in the following research question:

Research Question 3.1: Is there a difference in an indicator for monotony depending on the type of break activity?

However, a further question asks for additional strategies, which directly refer to the experience of air traffic controllers. As experts within their environment, controllers are asked for their recommendations:

Research Question 3.2: Which countermeasures can be defined that reduce the development of a state of monotony?

Through the definition of a set of applicable strategies a kind of tool-box should be offered for different groups which might be in the situation to deal with the issue of monotony.

3.3. RESEARCH ACTIVITIES

3.3.1. Combination of a Simulation and a Work Setting Approach

The definition of a set of research activities is required to answer the presented research questions. The following subchapter describes the methodology and discusses advantages and disadvantages of the selected procedures. Generally, two approaches can be applied to investigate the phenomena of monotony. Experimental studies in the laboratory allow for a full control of influencing variables and keep the internal validity high. Conversely, field studies are reduced in controllability but dispose of higher external and ecological validity.

To control the advantages and disadvantages of each, a combination of both approaches was chosen to arrive at a comprehensive description of monotony and was regarded as complementary:

- A simulation approach aimed to investigate the concept of monotony under a systematic variation of repetitiveness and traffic density in traffic characteristics to cover both facets of monotony. A small-scale experiment was conducted to validate the experimental approach applying the descriptive data analysis method according to Abt (1987) and define confirmative research hypotheses for the main study. With this background a main study was conducted to arrive at a conclusion for task factors.

- Since experiments do not represent the complexity of an operational environment, a field study was conducted to investigate the effects of selected factors from the simulation approach in the ops-room. This work setting approach was introduced because of several problems that are connected with simulation set-ups. Especially when the development of underload is investigated, there are several major risks. First, with continuous time on task the interest in the simulation might decrease and lead to risk seeking behaviors. Also, various aspects of the environment like the influence of colleagues cannot be simulated.

The discussion of the comparability of lab and field settings in psychophysiology was already led by Fahrenberg, Foerster, Schneider, Mueller, and Myrtek (1984), even though additional points come up when tackling practical research problems. A combination of basic and applied research does not necessarily contain a uni-dimensional research transfer from the laboratory to the field and may further include additional steps such as simulations or evaluations in restricted work settings. For example, in the current work the problem is posed by external facts and after linking theoretical concepts a simulation-based lab study was chosen. Skraaning (2003) discusses the problems of combining laboratory and field studies and sees the only solution in designing research settings as close as possible to reality, which is reached through conducting simulator studies in complex operational environments. Not fully applicable for the current work, it is noted, that ATC simulations are an important component in the formation and training of ATCOs, which justifies its application to derive valid conclusions. Also, an experienced sample helps to overcome the weaknesses of simulation settings.

Each of the objectives was addressed in the simulation set-up and in the field study. Different aspects were centered in each study approach to optimize the outcome, and studies built up on each other. The main objectives of the simulated approach were to describe how a state of monotony develops, how it can be measured and to determine the contribution of several factors in the controlled settings of a laboratory. This allowed including a number of variables that might not have been collected in a field study. It was assumed that both the uneventful and repetitive traffic conditions lead to monotony in an easy and uniform task. The focus was set on repetitiveness, which had not been investigated before. It was preferable to design a control condition, as it allowed the application of clearer manipulation criteria. The set-up also should not allow a classification within vigilance studies, even though related aspects were contained. As repetitiveness need not directly relate to task difficulty, a further independent factor was chosen. A preliminary experiment addressed questions of study design with the goal of defining the appropriate procedures and variables. From the following main experiment significant conclusions concerning the impacts of task characteristics, individual contributions and situational factors were deducted. The focus of the objectives in the field study was set to validate the results of the lab study in field conditions and to further elaborate on influences and strategies.

3.3.2. The Chosen Approach Towards Hypothesis Testing and Statistical Analysis

In the following chapter, some basic principles related to the approach chosen for hypothesis testing and statistical analysis will be discussed. The interested reader may find an introduction to conducting scientific experiments in Appendix D, as they lay down the basis for the current research activities and are especially dedicated to readers who want to have more information on issues related to experimental designs and statistical analysis.

One of the basic issues to be discussed in the following is related to alpha-inflation, as a high number of hypotheses are tested in the current studies. When several hypotheses are tested within one study, there is still the risk to reject a null hypothesis that is true because of alpha inflation, which means that the alpha error with an increasing number of hypothesis tests increases considerably. For example, with an alpha level of p=0.05, 5 % of the total number of hypotheses are rejected in case the null hypothesis is true. For this reason, depending on the amount of hypotheses considered for statistical decisions the alpha level is corrected. Several solutions were proposed how to solve this situation and are mentioned in Abt (1984). More recent approaches comprise the definition of a hypothesis space or the definition of an overall test parameter (Bauer, 1991, 1998). Loftus (1993) suggested to alternatively using graphical methods to get an idea about patterns and effect sizes. In addition, Sedlmeier (1996) discussed the use of confidence intervals and error-bar-plots, graphs and effect sizes especially in the context of explorative data analysis in addition to traditional procedures. He also demonstrated that often two approaches of hypotheses testing are mixed up, that is the Fisher significance test and the Neyman and Pearson approach. The importance of the latter is that it considers the expected effect sizes and defines potential risks of wrong decisions that influence the selection of alpha and beta errors as well as the sample size in a study, while the Fisher significance test defines a prior alpha level and the hypothesis is maintained or refused depending on the magnitude of the p-value. A further approach often chosen is the increase of the alpha-probability to 0.1, which does however not reduce the problem of insufficient beta and sample sizes. Finally, alpha correction has also been exposed to further critics (Perneger, 1999).

As Abt (1984) recognized, studies do not always have the scope to confirm differences between true effects (confirmatory analysis) which result in a final conclusion. Because of a perceived gap between the potentials of explorative data analysis and confirmative data analysis, Abt (1987) elaborated the descriptive data analysis (DDA) which is classified between these concepts. This will be further described, as it is the basis for the integration of the current studies in the framework. Exploratory studies try to define a totally new subject and are intended to generate new hypotheses. But they require testing these hypotheses in a new study, which is a disadvantage in cost intensive studies. DDA can be applied in studies where already some idea about the research subject exists, as for example in the current study, where the literature review allowed gaining some insights for most of the aspects of interest. In such cases, Abt (1987) proposed to formulate hypotheses on the assumptions already existing but not to use the confirmative hypotheses testing approach. He recommends the distinction of confirmatory hypotheses from descriptive hypothesis, where the latter "…yield hints at those of the N comparisons for which differences of true effects possibly exist" (Abt, 1984, p. 50). DDA does not require deriving any statistical decisions but gives a certain idea about the directions of the differences. Thus, it is possible to recognize "…patterns of descriptive significances associated with relevant effect differences if such appear to exist" (Abt, 1987, p. 81). In addition, a planned study may also include confirmative and descriptive parts, where the confirmative parts are established before starting the study. It is noted that there is no unified opinion towards this approach. For example Erdfelder (1994, p. 77) warned of using confirmatory statistical methods within exploratory data analysis. Statistical analysis procedures do aggregate data to statistics which might not reflect the occurrence of relevant variations in the raw data that thus contribute to clarify certain aspects in the subject of interest.

Because of the described strengths of Abt's approach (1984, 1987), this basis was chosen for the current research activities. The argument of Erdfelder is considered through sufficient consideration of individual data when describing the results. The small-scale experiment (Study I) is undertaken in line with the propositions by the descriptive data analysis approach. The simulated lab experiment (Study II) and the work setting approach (Study III) are seen as confirmatory studies. Both do however contain additional descriptive hypotheses. For this reason, some of the research questions were transferred in primary (main) hypotheses, while others were considered in secondary (additional) hypotheses. The alpha level is corrected for the main confirmative hypothesis to counteract the problem of alpha inflation. Additional descriptive hypotheses are marked but do not result in final conclusions.

3.4. SUMMARY

A range of problems was determined that limit the understanding of monotony in ATC. For this reason, objectives were defined which addressed the definition of task factors as well as the description of monotony and other critical states. Finally, countermeasures applicable in the field as well as in the future concept development need to be regarded. The research activities asked for an integration of simulation and real-time settings and considered confirmative and descriptive analysis approaches.

4. STUDY I: A SMALL-SCALE INVESTIGATION OF FACTORS RELEVANT TO DESCRIBE MONOTONY IN SIMULATED ATC

4.1. OBJECTIVES AND HYPOTHESES

The first objective introduced in chapter 3 addresses the definition of factors that evoke a state of monotony. These candidate factors can be derived from the review of research in related areas. Because they have never been systematically investigated in simulated ATC, a small-scale experiment was conducted to assess these factors and their contribution to monotony. According to the second objective, a further scope was to define indicators that characterize a state of monotony. The manipulated task characteristics were repetitiveness and also low and high levels of dynamic density, where the latter was seen as a form of uneventfulness. Psychophysiological assessment was employed to assess the ATCO's state. In addition, it was evaluated if the planned procedure was appropriate to elicit a state of monotony as far as it concerns the duration and the design of the traffic scenarios. Especially, the selection of appropriate materials as well as the definition of time intervals for the administration of rating scales and analysis of physiological indicators was seen as crucial. Thus, the following research hypothesis summarized expected effects on different psychophysiological indicators depending on task factors:

Hypothesis I.1$_D$[5]: There is a difference in physiological, subjective and performance indicators in different measurement periods depending on repetitiveness and sequence of dynamic density.

The applied procedure for the statistical hypotheses testing and the statistical analysis is further described in 4.2.6.

4.2. METHOD

4.2.1. Experimental Design

The experimental design presented in Table 5 consisted of three (respectively four in selected indicators) independent variables, resulting in a 2 x 4 x 2 (x 3/16) mixed design

Table 5: Experimental design of study I

		REPETITIVENESS								
		repetitive					non repetitive			
Run 1	DD	l	h	l	h		l	h	l	h
	Interval	1 2 3	1 2 3	1 2 3	1 2 3		1 2 3	1 2 3	1 2 3	1 2 3
Run 2	DD	l	h	l	h		l	h	l	h
	Interval	1 2 3	1 2 3	1 2 3	1 2 3		1 2 3	1 2 3	1 2 3	1 2 3
	n	1	1	1	1		1	1	1	1
	n	4					4			
Note. N=8. DD=Dynamic Density: low (l) vs. high (h); Interval (i)=Interval during run is included for i=3 measurements, but was also varied with i=16.										

5 marks a hypothesis according to the descriptive data analysis (DDA)

The between-subject factors were *repetitiveness,* varied between repetitive and non repetitive traffic situations, and *sequence of dynamic density (DD),* varied between low and high and completely permutated for two scenarios. Their operationalization is described in the following subsections. The within-subject variables were *run* (first/second scenario) and the *intervals during run* (3 respectively 16 intervals during each scenario). Participants were randomly assigned to the experimental conditions.

4.2.1.1. Independent variables (IV)

<u>Repetitiveness</u>

The condition of repetitiveness in the simulation scenarios was varied between repetitive versus non repetitive traffic situations. According to the diversity in available definitions for monotony an application of the term *repetitiveness* was preferred that directly referred to the task. In terms of Cox (1985) repetitive means a predefined cycle of elements is occurring several times in a sequence. But this definition is not including the required action to deal with these elements. Therefore, it was extended for its application in ATC, based on the assumption that a presentation of task elements relates to an action which is required from the controller to reach a task goal. However, a controller has many ways to deal with a situation. For this reason, to result in monotony, a restriction in possible solutions is indispensable. Additionally, *predefined* solutions can hardly be integrated in a definition for ATC as even repetitively planned flights might frequently result in deviations and unique situations. But, as was described in the introduction, controllers build up a mental picture that contains strategies to deal with a situation. For this reason, it can be assumed that similar strategies are continuously repeated in the work environment, once they were successful to gain routine and avoid increased workload.

The repetitiveness in the simulation scenarios was implemented through the manipulation of potential conflict situations throughout a virtual sector and is supported by the following arguments:

- ATCOs anecdotally state that in everyday life aircraft often meet at the same critical crossing points;
- ATCOs remain busy through continuously scanning and updating their mental picture;
- In certain situations, such as restrictions in neighboring sectors, the appearance of potential conflicts can hardly be prevented despite available flight information; and
- Potential conflicts require an action to avoid separation infringements.

There are a number of possibilities how to implement repetitiveness. For example, a focus could be set on an alternative area of control. Approach control is commonly seen as a highly repetitive task and controllers are more restricted in their actions when sequencing aircraft for landing. Nonetheless, because of often related time pressure particularly at bigger airports, it could not be predominantly considered as evoking a state of monotony in the theoretical sense, but might be rather related to stress. In contrast, enroute control disposes of a higher number of potential elements that might be defined as repetitive, such as the traffic routes, the flight plans of the aircraft, the crossing points, etc. Hence, an emphasis on enroute control was preferred for the investigation. As a final point, differences in the responsibilities between planner and executive controllers favor a focus on the executive position.

The scenarios included potential conflicts in constant 3-minute-intervals. Each conflict would have resulted in a very close near-miss without the controller taking appropriate action. Sufficient time to recognize the conflict was guaranteed. In repetitive traffic scenarios, participants were presented with equal potential conflict situations at the same crossing point labeled RINAX (see Figure 1). The constant situation consisted of an aircraft in departure meeting an incoming northbound aircraft in 3-minute-intervals. The aircraft requested to climb to its cruising level, usually at flight level FL310 or FL330, and hand over procedures required to exit the sector at these levels.

In case of a potential conflict, controllers were instructed to level off the aircraft in departure at FL290. In the repetitive condition, the incoming aircraft were at FL300. As such, if the departure would not have been leveled off in time, a Short Term Conflict Alert (STCA) would have activated. In the non repetitive condition, the potential conflicts involved different aircraft at varying crossing points throughout the sector in the same 3-minute-intervals but without a potential conflict between the aircraft passing RINAX.

The experimental situation modeling the repetitiveness thus met the required criteria for an eventual state of monotony mentioned by Bartenwerfer (1960) and Smith (1955): the sameness of the work process, the maintenance of a constant attentional focus, the low difficulty of the task and the unchanging work environment.

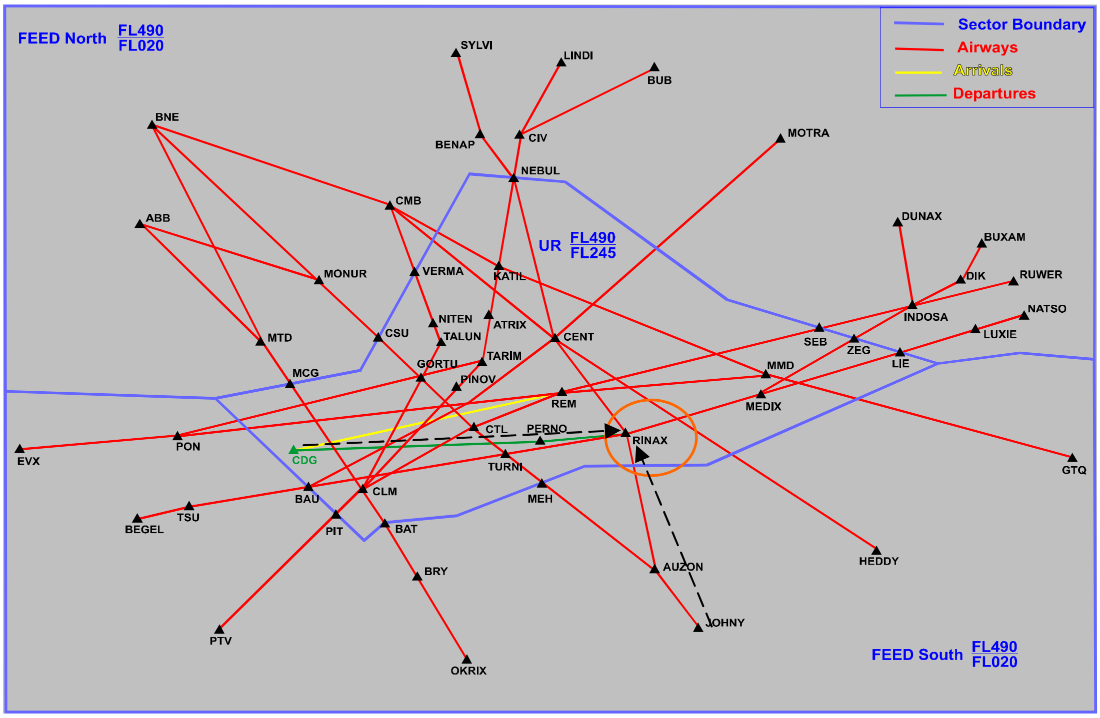


Figure 1: Sector map indicating potential conflict situation at RINAX between eastbound departing traffic and northbound traffic.

Sequence of Dynamic Density (DD)

The sequence of Dynamic Density (DD) was varied between high and low and helped to determine any interactions between task load and repetitiveness within their effect on monotony. This matter is not only implicitly related to uneventfulness, but also to complexity. As known from previous studies, the processing of information changes with increased complexity. Under higher traffic load, the way of information request becomes more constant and is also reflected in more uniform controller strategies in restricted airspace (Colterier, 1971, in Mogford, 1995). When few possible solutions are available, planning has to be done at an early stage, which has implications in sectors with high traffic density and restricted maneuvering space. This led Sperandio (1978, in Mogford, 1995) to the conclusion that controllers select operating procedures based on economy and use only the most critical data in high traffic density.

Generally, the number of aircraft is used as a thumb rule to predict the workload in a sector. But the number of aircraft does not adequately reflect the difficulty of the work situation, which appeared to be better expressed in *complexity*. Mogford, Guttman, Morrow, and Kopardekar (1995) reported a variety of factors contributing to the complexity of the situation, such as the traffic mixture, the number of crossing points, the number of climbing, descending traffic, *etc*.

This approach did however include the shortcomings of a static concept. The consequently developed concept of Dynamic Density (DD) considered continuous changes in these factors and allowed a better description of how a traffic situation developed over time (e.g., Laudeman, Shelden, Branstrom & Brasil, 1998; Kopardekar & Magyarits, 2003; Masalonis, Callaham, & Wanke, 2003). In this regard, it is the collective effect of all factors that contribute to the sector level complexity or difficulty at any given time. To demonstrate a similar idea of merging complexity and dynamics, Vogt, Adolph, Ayan, Udovic and Kastner (2002, p. 357) used the term *dynaxity*. While the concept of DD did not categorize factors, in this model *complexity* represented rather the spatial structure of a number of elements and *dynamics* reflected changes of the features over time. However, DD was preferred as it allowed the control of the variation in traffic in the course of the scenario.

The concept of DD was adapted to the needs with a focus on the most important factors determined in Laudeman et al. (1998). Within a certain range, in units lasting for three minutes the number of aircraft, heading changes, changes and predicted conflicts were kept constant. The average number of aircraft was 57 per hour and represents moderate traffic load. This kept controllers busy through continuously checking traffic. The manipulation between high (h) and low (l) DD was implemented with additionally required level changes of aircraft in the high DD condition. As shown in the studies of Kopardekar and Magyarits (2003), the level changes are one of the most important components in formulas defining dynamic density. To distinguish between the two treatments, a difference of greater than 1 SD was selected for the DD measure (see Appendix A.1.1 for computation details). A greater difference in traffic samples might have led to highly busy conditions and thus evoke intense stress reactions interfering with monotony. Because of their impact in the current study, the weighting-factor on the level changes was increased by 2. Finally, the traffic samples were tested by two operational experts who stated sufficient realism.

Although DD was repeated within subjects, the sequence of DD (l-l, h-h, h-l, l-h) was included as a between-variable. As it was not clear, which combination of the DD manipulations might have been most effective, in the preliminary study all possible combinations were built-in. The advantage of this procedure is a more precise estimation of effect sizes.

Run

Participants were presented with two scenarios of 50 minutes each. The changes in the dependent variables were compared between the first and the second run.

Interval during run

This factor was included in the analysis of subjective ratings that were collected three times during each run. Also, physiological measures were analyzed in 3-minute-intervals throughout each run, resulting in a total of 16 intervals.

4.2.1.2. Dependent variables (DV)

The assessment of task effects on individuals is based on the multi-level measurement approach which considers physiological variables, subjective ratings and behavioral indicators such as performance.

Physiological measures

Table 6 shows a summary of physiological measures. Physiological indicators are well known to significantly vary between and within individuals due to the psychophysiological principles described in chapter 2. Hence, multiple indicators are recommended to develop a more complete understanding on task-specific reaction patterns. The selected 3-minute-intervals for describing changes in physiological measures correspond with the manipulations of the traffic-characteristics and have been found as sufficiently fine-grained in earlier studies to reflect variations in task demands. Even though heart period (Inter-Beat-Interval, also called IBI = 1/BPM) is known for better distribution properties, heart rate was analyzed, since the more commonly used beats per minute (bpm) demonstrate higher face validity.

Table 6: Summary of physiological variables (Study I)

INDICATOR	DEPENDENT VARIABLE
ECG	Average heart rate (corrected for baseline) in 3-minute-intervals during run
	Average heart rate (corrected for baseline) in performance tests
	Average heart rate variability in 3-minute-intervals during run
	Average heart rate variability in performance tests
EDA	Average skin conductance level (corrected) in 3-minute-intervals during run
	Average skin conductance level (corrected) in performance tests
EOG	Average number of blinks in 3-minute-intervals during run

Subjective measures

The psychological assessment addressed individual reactions related to cognitive, emotional, energetical, and motivational aspects. For this reason, a variety of scales and questionnaires during and after the scenarios were administered. The scales contained items for the individual perception of aspects during the scenarios and elicited critical states, mood, workload and situation awareness. Table 7 shows a summary of the materials which are described in detail in 4.2.4.

Table 7: Summary of subjective variables and applied scales (Study I)

LEVEL	DEPENDENT VARIABLE
Cognitive, emotional and motivational indicators (TSI)	Average scores of *attentiveness* at 3 points of measurement during run Average scores of *fatigue* at 3 measurement points during run Average scores of *boredom* at 3 measurement points during run Average scores of *irritation* at 3 measurement points during run Average scores of *strain* at 3 measurement points during run Average scores of *concentration* at 3 measurement points during run Average scores of *motivation* at 3 measurement points during run Average scores of *sleepiness* at 3 measurement points during run
	Average scores in *feeling of monotony* after each run
Mood (UWIST)	Average scores in *hedonic tone* after each run Average scores in *tense arousal* after each run Average scores in *energetic arousal* after each run
Workload (NASA-TLX)	Average scores in *mental demand* after each run Average scores in *physical demand* after each run Average scores in *temporal demand* after each run Average scores in *frustration* after each run Average scores in *effort* after each run Average scores in *performance* after each run Average scores in *overall workload* after each run

LEVEL	DEPENDENT VARIABLE
Critical States (SOF)	Average scores in *stress* after each run Average scores in *fatigue* after each run Average scores in *monotony* after each run Average scores in *satiation* after each run
Situation awareness (SASHA)	Average scores in 8 items after each run

Two types of variables were used to assess monotony, the item of Thackray et al. (1975) to rate the subjective feeling of monotony and the Scale of Feelings (SOF) subscale monotony, which consisted of a combination of items describing the phenomena of monotony.

Performance measures

The difficulty of performance assessment in ATC was generally discussed by Manning (2005). Two approaches were selected to evaluate the performance (Table 8). The primary task indicators were deducted from the scenario log-files and examined in terms of Short Term Conflict Alerts (STCA). Performance tests were introduced after both scenarios to assess after-task effects on reaction time, concentration and anticipation capabilities. The frequently applied secondary task technique was not considered an appropriate alternative as it might have influenced the state of monotony and made the task more interesting. It was expected that differences in the experimental conditions were demonstrated through consecutive performance tests, similar to the study of Schellekens et al. (2000). The advantage of this procedure was that it did not directly interfere with the primary task but nonetheless demonstrated if cognitive functions were impaired differently depending on the varied factors after a certain time on task. Hidden costs of task execution may be demonstrated when the maintenance of task performance is required after the completion of the primary task. Performance degradation in that case was amongst others confirmed by Hockey (2003, p. 18) who explained that after effects would reflect a central state of fatigue. Thus, this approach allowed to determine task effects on different cognitive domains and supported the distinction between fatigue and monotony. This assumption was based on Bartenwerfer's results (1957), who found improved performance when the participants had the opportunity to change the task.

Table 8: Summary of performance measures (Study I)

LEVEL	DEPENDENT VARIABLE
Primary task performance	No. of Short Term Conflict Alert (STCA)
After-task performance	Vienna Reaction Test: Average scores in motor time Vienna Reaction Test: Average scores in reaction time Cognitrone: Average scores in reaction time to correct answers Time-Movement-Anticipation Test: Median of total deviation time in sec Time-Movement-Anticipation Test: Median of total direction deviation in pixels

4.2.1.3. Moderator and Control Variables

The influence of several nuisance variables was considered. The time of day effect was held constant. In the first study, state and trait variables (Table 9) were mainly collected to describe the sample and differences between the treatment groups, as the sample size was too small to be considered in an analysis of covariance or blocking designs. The preceding state was expected to have an impact of how likely someone experiences monotony. If someone is already fatigued at the onset of a task, one should perceive fatigue or monotony earlier during task execution and invest less effort to cope with repetitive situations. Therefore, the assessment of states at the beginning of the experiment contained important information.

Table 9: Summary of potential moderator variables and applied scales (Study I)

LEVEL	VARIABLE
Initial states of recovery and strain (RESTQ)	Average scores in aggregated *recovery* subscales Average scores in aggregated *stress* subscales Average scores in subscales (*General Stress, Emotional Stress, Social Strain, Conflicts, Overfatigue, Lack of Energy, Somatic Complaints, Success, Social Recovery, Somatic Recovery, General Recovery, Recovery Sleep*)
Action control style (ACS)	Average scores in *Decision-related Action Orientation* (AOD) Average scores in *Action Orientation after Failure* (AOF) Average scores in *Action Orientation during Successful Performance* (AOP)
Morningness-eveningness-preference (MEQ)	Average scores in *morningness-eveningness-preference*
Boredom proneness (BPS)	Average scores in *boredom proneness*

Age, expertise, gender, vision, and handedness were collected to describe the sample. Body mass index was calculated from weight (in kg) and height (in cm). During the experimental session the time, room temperature, body movements, respiration, and further information on the initial state of the subjects were recorded. Due to the used interval length, there was however no need to correct heart rate variability for respiration influences.

4.2.2. Procedure

The experiment was run between 27th April and 11th May 2004 in the Human Factors Lab at the EUROCONTROL Experimental Centre (EEC) in France. For pre-information and preparation purposes volunteering participants received a controller handbook that contained information and instructions for the set-up, the planned procedure and the provided questionnaires (available in an electronic appendix). Subjects were informed that the scope of the experiment was to understand what makes the task of an air traffic controller interesting. It was avoided to talk about monotony. The handbook also contained a biographic form and various questionnaires for trait factors that were filled in by participants prior to arrival. If they agreed to participate in the investigation, they were asked to complete the study consent form before they came to the experiment. In addition, participants were also asked to refrain from eating and drinking coffee during the experimental session. The following picture (Figure 2) shows the experimental set-up with one of the participants.

Figure 2: Demonstration of the experimental set-up

The experimental session included various phases (Table 10). The introductory part started with a briefing, using a standardized set of instructions. Participants were given the opportunity to ask any questions pertaining to the study, prior to the application of the electrodes for physiological measurements. It is recommended to attach electrodes at least 20 minutes before starting recordings to improve conductivity. Before attaching the electrodes, skin was cleaned with alcohol respectively water for the measurement of electrodermal activity (EDA) and prepared with electrode cream. Subjects were also instructed to avoid extensive moving during the trials and scenarios. A practice session allowed the participants to familiarize themselves with the simulation set-up. An exercise scenario of 30 minutes was provided to get used to the early demonstration and evaluation platform (eDEP; Schaefer & Smith, 2006) and to test the Vienna Test System (VTS).

Table 10: Experimental procedure (Study I)

TIME	STEP	TIME in min	TIME tot in min
14.00	**BRIEFING & FAMILIARIZATION**		
	Welcome and summary of controller handbook	10	10
	Attach electrodes and answering questionnaires in the following order: ECG, respiration band, movement sensor, EEG, EOG, EDA, SPO2 Questionnaires for Initial State, RESTQ	40	50
	Training on simulator and Vienna Test System (VTS)	40	90
15.30	**SCENARIOS**		
	Baseline 1+ TSI	3	93
	Scenario1	50	143
	Baseline+TSI+ UWIST+NASA+SOF+SASHA+Reconstruction Interview	10	153
	Rest break	5	158
	Baseline+ TSI	3	161
	Scenario2	50	211
	Baseline+TSI+ UWIST+NASA+SOF+SASHA+Reconstruction Interview	10	221
	Performance Tests (VTS)	20	241
18:01	**AFTER SCENARIOS**		
	Remove electrodes	10	251
	Debriefing	20	271
18:31	**END**		**271**

After familiarization with the environment, participants performed two traffic scenarios that contained repetitive or non repetitive traffic and low or high DD. Each of the scenarios lasted 50 minutes after ATCOs took the scenario under control, which had been advanced for 10 minutes. Preceding and following each run, 3-minute-baseline recordings were taken in a relaxed resting condition with closed eyes. It is noted that the instruction was given to work according to ICAO standards. Physiological measures were collected with a dedicated recorder throughout the runs. The items based on Thackray et al. (1975) were filled in before each run and after 20, 35 and 50 minutes in the scenario. The participants were instructed to accommodate the questionnaire as soon as the tasks allowed a short interruption. NASA-TLX, UWIST and SOF were administered after each scenario. A post-interview after each run asked for special occurrences during the scenario and a debriefing concluded the session.

During the session an experimental checklist was kept by the experimenter to record control variables and follow the completion of single steps. An average experimental session lasted approximately 4 hours 18 minutes (SD=13 min). Videos of the scenarios were recorded to collect behavioral indicators that eventually allowed the interpretation of results, but were however not planned to be submitted to further analysis.

4.2.3. Participants

Volunteering participants were recruited amongst the operational experts placed at EEC who had been licensed as ATCOs. A brief email was sent to potential participants who were informed about the experiment and its procedure. A sample size of n=8 was considered sufficient in this small-scale experiment to estimate the size of the effects that can be expected in the main study, accepting that significant results would probably not be found due to low power. The factor repetitiveness was expected to demonstrate differences in the subjective ratings for monotony and physiological indicators. It was equally important to include a combination of all manipulations of DD at least once, as a decision about the sequence for the main experiment needed to be taken.

Table 11: Descriptive statistics for biographic and state and trait variables as a function of repetitiveness

VARIABLE	REPETITIVENESS	M	SD
Action orientation after failure (AOF)	non repetitive	6.50	2.65
	repetitive	7.25	3.77
Decision-related action orientation (AOD)	non repetitive	7.50	3.70
	repetitive	10.25	1.26
Action orientation during successful performance (AOP)	non repetitive	8.25	2.36
	repetitive	9.25	1.50
State of recovery(*)	non repetitive	2.74	0.57
	repetitive	3.56	0.78
State of strain	non repetitive	1.10	0.32
	repetitive	1.29	1.21
Boredom proneness score	non repetitive	95.25	20.71
	repetitive	94.00	4.24
Morningness-eveningness score	non repetitive	49.00	10.09
	repetitive	52.00	6.32
Age in years	non repetitive	45.25	7.93
	repetitive	49.75	6.44
License in years[a]	non repetitive	14.67	5.03
	repetitive	13.25	5.12

Note. N=8. [a]N=7. ***p<.001. **p<.01. *p<.05. (*)p<0.2.

The sample group consisted of seven male and one female operational experts aged between 37 and 56 years (M=47.5, SD=7.1). They were of four nationalities (British: n=4; Dutch, Austrian, Belgian, French: n=1) and had been fully licensed as ATCOs between 6 and 20 years (M=13.9, SD=4.7) in European control centers. Only one of them did not have instructor experience and two did not participate in previous simulations at EEC that used the eDEP environment. All participants were in good health with an average Body-Mass-Index of 24.4 (SD=1.9). The vision of three operational experts was corrected to normal; one of the participants was left handed.

Three operational experts had ratings for tower/approach/enroute, two of them had been rated for enroute and one each had ratings for approach, approach/tower or approach/enroute. Further variables were assumed to influence the interpretation. The descriptive statistics were described for experimental groups with the main focus on the repetitiveness factor (Table 11). Statistical analysis revealed a tendencially significant effect of repetitiveness in the aggregated recovery subscales ($t=-1.69$, $df=6$, $p=.142$). No difference in indicators were found in the groups assigned to the experimental DD conditions (all $p>.209$).

4.2.4. Material and Apparatus

4.2.4.1. Simulation environment

The controller working position (CWP) included a 28"LCD monitor with keyboard and mouse for inputs; Short Term Conflict Alert (STCA) was available and Reduced Vertical Separation Minimum (RVSM) for Europe applicable. Participants worked on a standalone sector with two automatic feed sectors. To avoid social and communication influences, runs were conducted individually with the controllers taking over executive and planning roles. Pseudo-pilots were not included. The scenarios were run on the eDEP platform developed by EUROCONTROL. The experiment was based on four traffic scenarios. The semi-generic upper airspace created for this experiment (FL 250 – FL 600) involved a sector with arriving and departing traffic from a major airport. As Guttman and Stein (1997, quoted in Manning & Stein, 2005) found, using a generic airspace can be expected to have no impact on the results. The simulation environment allowed recording controller inputs and the traffic evolution on a log-file.

4.2.4.2. Vitaport

Physiological recording was accomplished using Vitaport 3 of Temec Instruments, NL (Jain, Martens, Mutz, Weiß & Stephan, 1996). This device can be used for a wide range of applications and consists of a 12-bit analog to digital converter. Signals of varying sampling frequencies are read through a separate channel, pre-processed and stored on a PC-card. Settings used for sampling and recording are presented in Table A-5. Online-viewing of the recording process is possible, and the samples of all channels are stored in one file. Ag/AgCl electrodes were employed for ECG, EEG, EOG and EDA. The ECG is recorded from three electrodes (of which one ground) placed on the subject's chest. The EEG electrode was fixed at the position CZ with a reference placed at the right mastoid and filled with Quick-Gel electrode jelly (Med-Suppliers, NL). Recordings were made with silver electrodes applied to the skin with collodion and resistances kept below 5 kOhm. For horizontal eye movements two electrodes were fixed with adhesive rings at the left and right eye below the eye brows close to the canthus of each eye, and for vertical movements two electrodes were positioned at the upper and lower side of the eye before filling the cups with conductive cream (Quick-Gel, Med-Suppliers, NL). In agreement with the recommendations in Boucsein (1988), EDA electrodes were attached to the sole of the left foot and filled with an isotonic electrode cream (Biopac Inc.). The respiration was recorded using a strain gauge attached to a belt wrapped around the thorax. A sensor for movement recordings was placed at the chest and fixed with a tape. To collect peripheral pulses a sensor (Nonin Flex Sensor System) was placed at the left finger. It is based on the principle of reflection that depends on how much oxygen the blood contains; this data was however not further analyzed.

4.2.4.3. Vienna Test System (VTS)

VTS represents a computerized test battery (Schuhfried GmbH, AT), which contains selected tests for psychological diagnostics. In this experiment, especially the impairment of certain cognitive functions in consequence of task execution was of interest. The selected tests were demonstrated on a standard HP notebook and a response panel was used as an input device.

Vienna Reaction Test S5

In the Vienna Reaction Test (Schuhfried & Prieler, 2002), reaction time is measured through tasks containing single or multiple-choice answer formats. The general forms cover the areas of alertness and the ability to suppress an inadequate reaction. In version S5 the stimulus modalities ‘light’/’tone’ and the characteristics ‘red’/’yellow’ are available. The respondent is instructed to press the reaction key only when the relevant stimuli ‘yellow’ or ‘tone’ are presented. Through the use of a rest and a reaction key the measures can be divided into reaction time (time to release the rest button after stimulus presentation) and motor time (time between releasing the rest button and pressing the reaction button). Median and range (quartile) of reaction and motor time and the number of correct, omitted and incomplete responses are obtained. Cronbach’s Alpha[6] is r=.83 for reaction time and r=.94 for motor time and the administration lasts for 9 minutes. That reaction time is related to monotony was found in prior experiments. Thackray et al. (1975) found that the single longest response latencies were higher in the group that experienced boredom and monotony in simulated ATC. Bartenwerfer (1960) found no difference in reaction time before and after driving, but faster reactions after changing the task, which he explained with the importance of changing the task.

Cognitrone S6 (COG)

Cognitrone (Wagner & Karner, 2003) is used for the assessment of attention and concentration. It requires the comparison of an abstract figure with a sample and to judge if they are identical. Concentration is defined as the ability to direct attention to a task for a long time in order to attain a stable performance. In the introduced version S6 a participant has to compare a total of 308 complex figures with a model and answer within 1.8 seconds through pressing green and red buttons on the panel if the figures are identical or not. To guarantee a correct response, the respondent has to find a compromise between speed and accuracy of items. The number of correct and incorrect reactions with the corresponding median and range of reaction time are reported. These variables express the respondent’s ability to analyze patterns within a given time limit. Reliability for S6 is r=.93 for the sum of correct reactions and r=.97 for the sum of incorrect reactions. In the study these indicators were introduced to understand how repetitiveness and time-on-task affect the ability of ATCOs to concentrate.

Time-Movement-Anticipation S2 (ZBA)

The scope of the ZBA (Bauer, Guttmann, Leodolter & Leodolter, 2003) is to assess to which extent a person can project into motion as this is especially important for ATCOs. In Figure 3 it is demonstrated how a respondent has to estimate when and where a green dot disappearing at the first red line would have reached a second target line and press a button. The prior movements range from simple lines to curved paths and differently modulated sine curves. Registered scores include the error in time anticipation (measured as the time difference in seconds including 2 decimals from the correct response) and the error in motion anticipation (deviation from the correct position in pixels). Controllers executed the short form S2 that contains 12 items. This version lasts for about 10 minutes. Estimations for the internal consistency (Cronbach’s alpha) are only available for the long form and have been stated with r=.98 for median deviation time and r=.76 for the median deviation direction.

[6] A commonly used indicator of reliability: reflects how well a set of items measures a latent one-dimensional construct.

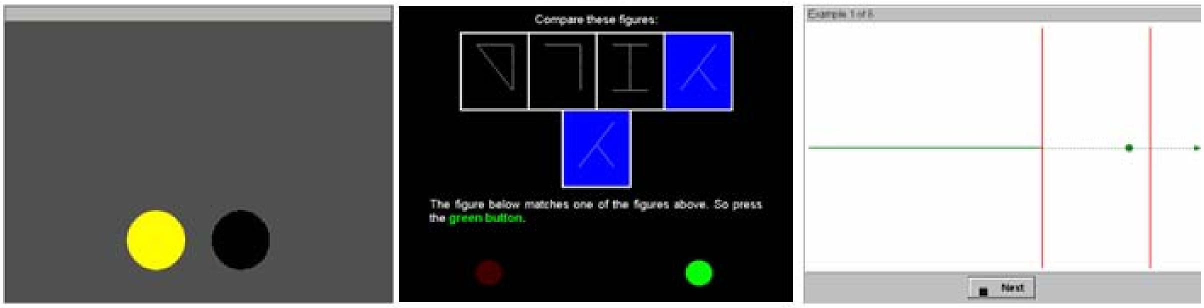


Figure 3: Screenshots of Reaction Test (a), Cognitrone (b) and Time-Movement-Anticipation (c)

4.2.4.4. Subjective Assessment

Biographic questionnaire

A biographic questionnaire was included in the controller handbook to collect major information of the participants, such as age, gender, nationality, native language, experience, ratings, handedness, vision, and recent simulation participation.

Questionnaire for Initial State

This questionnaire is based on the original form introduced by Janke (1976) and was updated for the investigation of energy drinks in an Austrian pilot sample (Deixelberger, Kallus, & Tischler, 2003). It addresses the general somatic initial state and was used as a questionnaire to better control influences that occurred before the experiment. Questions concern the activities during the day, last nights sleep, nutrition and alcohol, medicaments, and nicotine consumption.

Recovery-Stress-Questionnaire (RESTQ)

This questionnaire (Kallus, 1995) identifies different areas of stress and recovery states. It is based on the hypotheses that an accumulation of stress with insufficient opportunity for recovery leads to a compromised psychophysiological state. Questions address how often the person was exposed to stress situations over the last days and how often recovery activities were experienced. The current version exists of 48 items with 7 specific scales concerning stress and 5 specific scales for the areas of recovery, which are presented in Figure 3. The controllers were asked to rate the frequency of activities and how they felt within the last 3 days/nights on a 7-point rating scale ranging from ‘never’ (0) until ‘always’ (6). The psychometric qualities have been well-investigated and internal consistencies range between r=.80 and r=.97 with a median of r=.92.

Table 12: Overview and description of RESTQ subscales

	SUBTEST	ABBREVIATION	EXAMPLE ITEM
Stress	General Stress	genstr	I felt down
	Emotional Stress	emostr	I was in a bad mood
	Social Strain	socstr	I was angry with someone
	Conflicts	conflict	I felt under pressure
	Overfatigue	timepres	I was overtired
	Lack of Energy	noen	I was unable to concentrate well
	Somatic Complaints	somcompl	I felt uncomfortable
Recovery	Success	success	I finished important tasks
	Social Recovery	socrecov	I had a good time with my friends
	Somatic Recovery	somrecov	I felt at ease
	General Recovery	genrecov	I was in a good mood
	Recovery Sleep	sleep	I had a satisfying sleep

NASA TLX

The NASA Task Load Index (TLX, Hart & Staveland, 1988) was used to obtain a self reported assessment of mental workload during the experimental conditions. The original procedure is based on the weighted average of ratings after paired comparison between six subscales, described in Table 13: mental demand, physical demand, performance, temporal demand, effort, and frustration. This scale is based on a human centered framework, where workload emerges from the interaction between the requirement of a task, the circumstances under which it is performed and the skills, behaviors, and perceptions of the operator. Workload is defined as the “cost incurred by human operators to achieve a specific level of performance” (Gawron, 2000, p. 130). On a 20-point-scale ratings range between the end points labeled 0 (low) and 100 (high). From the weighted subscales an overall workload score is formed. Moroney, Biers and Eggemeier (1995) and Byers, Bittner, and Hill (1989) demonstrated that a simple summation on six subscales produced comparable means and standard deviations. They reported correlations between r=.96 and r=.98 with the paired comparison procedure. As the weightings do not significantly affect the resulting workload scores, in the current experiment an unweighted version was used. An overall index for the workload was obtained summing up subscales.

Table 13: Overview and description of NASA-TLX subscales

SUBTEST	ENDPOINTS	DESCRIPTION
Mental Demand	Low - High	How much mental and physical activity was required (thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, exacting or forgiving?
Physical Demand	Low - High	How much physical activity was required (e.g., pushing, pulling, turning, controlling, activating, etc.)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?
Temporal Demand	Low - High	How much time pressure did you feel due to the rate or pace at which the task or task elements occurred? Was the pace slow and leisurely or rapid and frantic?

SUBTEST	ENDPOINTS	DESCRIPTION
Performance	Good - Poor	How successful do you think you were in accomplishing the goals of the task set by the experimenter for yourself? How satisfied were you with your performance in accomplishing these goals?
Effort	Low - High	How hard did you have to work (mentally and physically) to accomplish your level of performance?
Frustration	Low - High	How insecure, discouraged, irritated, stressed, and annoyed versus secure, gratified, content, relaxed and complacent did you feel during the task?

Gawron summarized the studies which employed NASA TLX. Primarily used in aviation, its sensitivity was found sufficient. An example reported by Hancock, Williams, Manning, and Miyake (1995) found a high correlation between TLX score and difficulty in a simulated flight task. Battiste and Bortolussi (1988) reported a test-retest correlation for the overall workload of r=.77. Because of contradicting results in the vigilance research, where increased workload was reported in underloading monitoring tasks, this indicator was also assessed in the current study.

<u>Scale of Feelings (SOF-II)</u>

This scale is the English translation of the *Belastungs-Monotonie-Saettigunsskala* (BMS) II by Plath and Richter (1984) amongst others used in the work of Rockstuhl (2002). It is intended to measure four types of critical states (see Table 14) that emerge as a short-term consequence of task execution. Each subscale consists of 10 items representing different aspects of the states and two parallel versions are available. In the original versions items have a dichotomous response format and a person is asked to judge whether he or she associates to the statement in the item or not. The psychometric properties of reliability and validity were reported as sufficient. Currently a state and a trait version are available with response formats ranging from 1 to 4. Generally, the total score for each subscale is calculated by averaging the item-difficulty-parameters for those items that the respondent identified as applying to him. A critical point of this questionnaire is that the different states are correlating (Rockstuhl, 2002), apparently a consequence of the way the scale was developed. Also it was recommended to introduce it after an extended working period (Richter, Debitz, & Schultze, 2002). In the current version the item scores were averaged for each subscale.

Table 14: Overview and description of Scale of Feelings (SOF) subscales

STRAIN	DESCRIPTION
Mental fatigue	a state of exhaustion and tiredness that may arise after prolonged time spend on the task or because of increased task complexity
Monotony	a state of boredom and lack of interest that may arise because of too low demands
Satiation	a state of disinclined testiness that may arise because of lack in perceived meaningfulness of the task
Stress	a state of aroused-frightened tenseness, nervousness and concern about the ability to fulfill the demands.

UWIST Mood Adjective Checklist (UWIST)

UWIST (Matthews, Jones & Chamberlain, 1990) assesses affective aspects of work strain. The original version has four sub-scales that are each made up of a combination of positively and negatively loaded items. Three 8-item-scales measuring bipolar mood dimensions were included in the ongoing research while the anger scale with only positively loaded items was excluded. Tense arousal contrasts feelings of anxiety with calmness, energetic arousal opposes vigor and tiredness, and hedonic tone confronts contentment with depression. Table 15 enumerates the items for each subscale. Participants judge each item on a four-point-scale ranging from 'definitely' to 'definitely not'. Internal consistencies for the three scales range from r=.86 to r=.88 and the scale showed satisfactory predictive and discriminant validity described by Matthews and his colleagues (1990).

Table 15: Overview and description of UWIST mood assessment subscales

SUBSCALE	POSITIVE ITEMS	NEGATIVE ITEMS
Hedonic tone (HT)	happy	depressed
	cheerful	sad
	contended	sorry
	satisfied	dissatisfied
Tense arousal (TA)	anxious	calm
	jittery	restful
	tense	relaxed
	nervous	composed
Energetic arousal (EA)	active	unenterprising
	energetic	sluggish
	alert	tired
	vigorous	passive

Thackray Scale Inventory (TSI)

In the studies of Thackray et al. (1975) ratings of boredom, irritation, attentiveness, fatigue, strain, and monotony were included on a 9-point-scale. They closely approximated the items used in the studies of Barmack (1939a, 1939b, 1939d). The authors did not conduct any reliability studies based on any theoretical background, items can however be considered as sufficiently validated from the theoretical point of view behind this research and available average values and standard deviations of reported studies can be compared with the own results. Also, the introduction of brief items was the only way to get an impression of participants' state and their interpretation of the situation during the scenarios and to compare the outcome with previous results, such as the study of Frankenhaeuser (1971) or Johansson and Sanden (1989). The application of validated questionnaires would have been too time-consuming for an assessment during the scenarios. The items of Thackray et al. were completed with ratings of motivation and concentration, as deployed in the studies of Johansson et al. (1996). Motivation was an aspect connected to satiation in the early work by Barmack (1939c) or Berman (1939a), and was successfully used in ATC studies by Vogt and his colleagues (2002). Concentration was considered theoretically relevant, as it is related to the concept of effort, where the controller has to put effort and focus on his task. Sleepiness was introduced to complement ratings of the more complex aspect of fatigue. The item of monotony was excluded from ratings during the scenario and included as an item for subjective feeling of monotony in the TLX rating, as it might have been too obvious for participants to find out the scope of the scenario. Still, it has to be considered that the administration of scales might interfere with the task execution and as such interrupt any negative feelings related to monotony through a change in the task.

Morningness-Eveningness-Questionnaire (MEQ)

This self-assessment questionnaire of Horne and Ostberg (1976) determines the diurnal type of a person. It contains 19 items addressing sleeping and waking behaviors and daily rhythms for doing different activities. The questionnaire was validated with circadian peak times from core temperature. Answers use a forced choice indicating definite/moderate morning type and definite/moderate evening type; five questions were answered with crossings on a time scale. A recent validation of the questionnaire in French middle-aged workers (Taillard, Philip, Chastang, & Bioulac, 2004) determined new cut-off points. Diurnal preferences were assumed to play an important role to modify alertness especially at night. As Ognianova et al. (1998) reported, there was a significant negative correlation between morningness and alertness, ratings of sleepiness, distractibility, and irritability in the early night shift in 22 operators of a thermoelectric power station. Because of unclear typologies and a lack of further validations, the raw scores were used in the current study. Traditionally composed types were reported for sample description.

Action Control Style Questionnaire (ACS 90)

Three subscales described in Table 16 were developed to measure action control style which assesses the degree of dispositional action versus state orientation on the basis of several easily accessible phenomenal consequences that are postulated by the theory (Kuhl, 1994b).

Each scale consists of 12 items with two alternatives, where one alternative describes action oriented and the other one state oriented behavior. To obtain the scores for each scale, action oriented responses are summed up for each scale. State orientation is reflected in low values, action orientation in high values. The internal consistencies satisfy traditional standards concerning measurement properties.

Table 16: Description of action control style (ACS) subscales

ABBREV	POLES	DESCRIPTION	EXAMPLE ITEM	α
AOF	Action orientation subsequent to failure versus preoccupation	describes state-oriented preoccupation that occurs if the action component of an intention is ill-defined or degenerated and one does not know what to do to reach a goal	"After a failure I find myself thinking for a long time about how it could have happened"	.70
AOD	Prospective and decision-related action orientation versus hesitation	refers to the aspect if the subject or the relational component of an intention is degenerated and the actor does not know whether to identify with the intention	"If something must be done I begin doing it without hesitating"	.78
AOP	Action orientation during successful performance of activities (intrinsic orientation) vs. volatility	assesses the degree to which an individual is able to stay with a pleasurable activity once initiated	"When I read something I find interesting I will sit and read the article for a long time"	.74

Boredom Proneness Scale (BPS)

This scale assesses the tendency to experience boredom (Farmer & Sundberg, 1986) and originally contained 28 items answered in a true-false response format. Items include statements such as "It is easy for me to concentrate on my activities". Vodanovich and Kass (1990) revised the scale to a 7-point Likert format to allow more variability in the responses to each item. This and other psychometric measures of boredom have been reviewed by Vodanovich (2003). For the 7-point-response format of the BPS internal consistency estimates ranged between .79 and .84. For the original version test-retest-reliability after one week was r=.83, and internal reliability r=.79.

Also, boredom proneness was significantly correlated with depression, hopelessness, and loneliness (Farmer & Sundberg, 1986) and relationships with sensation seeking, self-actualization scores (Vodanovich & Kass, 1990), and negative affect (Vodanovich et al.,1991) were found. In a further investigation Vodanovich and Kass (1990) found five factors in a sample of American undergraduates; that is external stimulation, internal stimulation, affective responses, passage of time, and constraint.

Boredom proneness was identified as a relevant characteristic for several reasons. In a vigilance performance task applying a clock test for 60 minutes in 33 undergraduate female students Kass, Vodanovich, Stanny, and Taylor (2001) found that sensory efficiency in the first 10 minutes related negatively to scores on the BPS. Relationships to absenteeism, tenure, and job satisfaction were found (Kass, Vodanovich, & Callender, 2001) in a sample of 292 workers of a manufacturing plant. Sawin and Scerbo (1995) found a significant negative correlation (r=-.30) between the BPS score and the proportion of hits on a flicker-detection task in 60 adults.

SASHA

Low situation awareness is a frequently mentioned factor related to aviation incidents and according to Endsley (1999) low situation awareness dominant in underloaded conditions has already led to incidents. Situation awareness is the momentary understanding of the current situation and its implications (Tsang & Vidulich, 2002, p. 177) and Vidulich (1990) described it as a concept concerned with the quality of information apprehended by the operator. Thus, the relation to the current study is that an operator needs to have an accurate picture of the current situation. The often cited model of Endsley assumes a constitutive progress from perception to prediction, but does however not consider that based on experience and the mental model actions might be decided before the elements are perceived completely. Even though, discussions of this concept go beyond the scope of this thesis.

There are various methods that have been created to assess situation awareness, summarized in Jeannot, Kelly and Thompson (2003). Each of them has been associated with advantages and disadvantages. To overcome some of the related problems, EUROCONTROL developed a questionnaire to assess situation awareness in simulations, which is available in form of an online-expert and a self-rating version. As in the current experiment it was not possible to include additional experts, the self-rating-version was used. However, related critics need to address the lack of calculated test metrics to evaluate the reliability and validity of this measure. For this reason, its application is rather seen as a contribution to the assessment of psychometric characteristics.

Reconstruction interview

In the prestudy, a reconstruction interview was introduced after each scenario. One part of the questions was based on the guide for reconstruction interviews developed during the Integrated Task Analysis (ITA) to investigate mental processes in air traffic controllers (Kallus, Barbarino, & Van Damme, 1998). Its purpose is "… to elicit the reaction to and strategies for resolving difficult situations, and to address the problem of additional load arising from co-ordination, planning, loss of time, etc. " (p. 36). It combines features of the critical decision method because it addresses critical events from the previous work period with verbal report methods based on a reconstruction of the situation. Thus, it links the collected subjective, physiological, and behavioral data to the subjective interpretation of the participants.

The first question addressed the STCA alerts that eventually occurred. Participants were asked for their reaction, the plans and what they did to solve the situation. In case there was no STCA alert during the scenarios, it was immediately continued with asking for the potential last conflict situation, their plans and actions how to cope with this situation with reference to the environment.

The disadvantage is that this method depends on how well the retrieval of the situation works, even though high memory of ATCOs for passed events has been described several times (e.g., Vogt et al., 2002).

Debriefing guide

A debriefing concluded each session. The questions addressed the experiment in general, attributed goals, its procedures, the used questionnaires and scales and the physiological measurement. Additionally, the traffic scenarios were addressed to reveal differences between the two runs, how realistic it was perceived and what was different from reality. Another aspect addressed the strategies to cope with such situations in reality, which critical states arose and how participants noticed them.

4.2.5. Data Processing

Different processing procedures were applied for physiological, subjective, and performance indicators. After recording physiological data, they were graphically displayed within *Vitagraph* before exported to and analyzed with *Acknowledge 3.7*. (Biopac Inc). The determination of heart rates was conducted with *Acknowledge*, where manual artifact correction was implemented. Sternbach, Alexander, Rice, and Greenfield (1969) discussed several artifact correction procedures. The selected one replaced the period of the artifact with the time period preceding the artifact. This approach was considered superior to excluding the whole interval because of its efficiency and also sufficient since only longer periods were analyzed and high temporal resolution was not required. Apart from that, artifacts rarely occurred. Three-minute-intervals were selected for the aggregation of the data. This was also the preferred interval length discussed by Sternbach et al. (1969). The comparison of ECG-correction procedures offered in *Vitagraph* and *Acknowledge* revealed that results obtained with ACQ were more reliable. Statistical analysis after using Vitagraph resulted in favorable results compared to *Acknowledge* (see Appendix A.1.2). This might be a consequence of the implemented automatic correction algorithm for undetected or missing heart rates which might have caused incorrect replacements. Acknowledge requires visual checks for artifacts and manual correction, while Vitagraph implies automatic processing. In consequence, *Acknowledge* was preferred to execute data processing even though higher investment of time was necessary.

All indicators were determined for 3-minute-intervals. Because of easy computation and its successful application in the study of Thackray et al. (1975), heart rate variability was chosen as the indicator to reflect heart rate fluctuation. In line with Walschburger (1976) the variance of the heart rate was preferred to the frequently used standard deviation as the square reinforces eventual effects. Also, the HRV was measured in the same 3-minute-intervals to compare other physiological measures and thus deviates from the commonly used 5-minute-intervals for SDNN. Additionally, the number of eye blinks and the skin conductance level were determined with *Acknowledge*. The EOG has a frequency range between 0.1 and 38 Hz and typically is below 20Hz. To remove the DC component but preserve the AC signal, a first order low pass filter with a 1-second-time constant (0.159 Hz) was used. The cut-off frequency of 38 Hz was used to reduce noise contamination and minimize unwanted EMG and EEG interference. Blinks were defined as peaks which reached a certain level that was individually determined. The electrodermal activity was processed as recommended by Walschburger (1975).

Finally, the summarized indicators such as mean and standard deviation of the processed indicators were stored as text-files and imported into SPSS 11. Subjective data were entered in a text editor, imported in SPSS 11 and integrated in an overall data file used for statistical analysis. Primary performance data was obtained through the recording of a log-file during the scenario execution and contained detailed traffic information. For the analysis of the small-scale study only STCA alerts were considered after the occurrence of a real STCA alert had been verified in the reconstruction interview.

Even though missing values were rare, the treatment of missing data was executed in the following way. Because of the low sample size the exclusion of a case was undesirable and authors recommend different approaches for an appropriate substitution of missing values (e.g., Bingham, Stemmler, Petersen, & Graber, 1998). As a general rule, missing data was interpolated along with the best fitted trend component. Due to system crash, this concerned the interval 16 in the first run of participant 7. Participant 3 lost the contact of the EDA electrodes from the beginning of interval 7 in the second run. In the performance indicators, system problems of the Vienna Test System resulted in missing data for the COG (n=3) and the ZBA (n=1).

Another important question concerns how to deal with the initial value (Kallus, 1992; Gratton, 2001, p. 919). The impact of the activation level has been intensely discussed and different solutions were proposed to deal with it. As recommended by Stemmler (2001, p. 25), the reactivity measure shall be chosen depending on the research question, since conclusions can be influenced by the selection of the reactivity measure. One of the possibilities to define a reactivity measure is to calculate the difference between a baseline and the experimental condition. This requires a discussion of the type of baseline. The initial baseline has the disadvantage that it might be influenced by effects of the experimental setup (e.g., excitement at the beginning). Jennings, Kamarck, Stewart, Eddy, and Johnson (1992) recommended 10 minute baseline measures before, during and after the experimental conditions, as stability was found to be high for heart rate and blood pressure. As it was not possible to determine the absolute night minimum of the baseline for an operator, as proposed by Fahrenberg et al. (1985, quoted in Stemmler, 2001), the approach of Jennings and colleagues was applied, where several baselines were collected during the experimental procedure. Thus, the frequently applied definition of a baseline as the information gathered at the beginning of a study from which variations found in the study are measured needs to be extended to a definition which considers a known value or quantity with which an unknown is compared when measured or assessed. Finally, due to temporal restrictions, baselines were recorded for 3-minute-intervals, which were found sufficient also in prior studies. An alternative correction, the level-correction compares the values of the experimental condition with the total number of measures collected for a person, including baseline and experimental conditions. In that case, the experimental effect would remain after consideration of time-on-task, the law of initial value and individual scale level preferences. This method is however just appropriate, if the number of baseline measures is not highly different from the number of experimental conditions. In that case, eventually available treatment effects would disappear. The correction method applied for the heart rate (HR) was the correction with averaged baseline-values (4 baselines collected). This procedure was not appropriate for the HRV, as it was generally fluctuating strongly and thus an averaged baseline would have confused the interpretation of the results. For this reason uncorrected values were used. Skin conductance level (SCL) was corrected after the procedure proposed by Lykken, Rose, Luther, and Maley (1966). This is however different for subjective ratings, where the variation in preferences towards lower or higher end points in scales varies more from the beginning. In the mentioned studies of Frankenhäuser (1971b) and Johansson et al. (1996) this issue was resolved through the rating of a standard situation. The current research applied the level-correction procedure for subjective measures, as the number of intervals was rather low compared to the physiological measures.

4.2.6. Statistical Analysis and Hypothesis Testing

The initially formulated research question addressed the difference in the development of multiple indicators over time depending on the traffic characteristics repetitiveness and dynamic density. As the scope of this study was to understand the development in different dependent variables, the statistical analysis was conducted according to the Descriptive Data Analysis (DDA), an approach suitable for multivariate analysis.

To determine the main effects of the independent variables, the statistical hypothesis for each indicator was formulated as a null hypothesis and generally stated that the mean parameters in the population do not differ in any of the conditions. As a factorial design was used, this hypothesis is tested for the effect of each between- and within-subjects factor. In addition, all higher-order interactions between these factors were investigated.

The indicators were submitted to analysis of variance (ANOVA), based on the assumptions of the general linear model. *Repetitiveness* and *sequence of DD* were included as between-subject factors; *run* and *interval during run* were within-subject factors. In contrary to the relevance of the repetitiveness factor, which was determined through a significant main effect, the determination of the statistical relevance of low or high DD was derived through the evaluation of the trend component in the significant DD x Run interaction. A significant effect indicated a different impact of high or low DD which could be further described to understand the direction of the effect. In the context of the DDA, assumptions for normal distributions and equality of variances were neglected. As it is known to have greater power (Stevens, 1999), univariate analysis was employed. As the results are used to detect any systematic patterns in the conditions, uncorrected degrees of freedom (df) were used. For the DDA, no post-hoc tests were conducted, as it was more important to analyze effects in the trend components. Descriptive significances demonstrate the probability to falsely reject the null hypothesis even though it is true. Exact p-values are marked as significant (p<.05*; p<.01**, p<.001***) or tendencially significant (p<.10[(*)]). In some instances, p-values between 0.1 and 0.2 will be included in the description of the results[7]. Linear and higher order trend components were indicated if the two-tailed probabilities were p<.05.

4.3. RESULTS

In the following sections, results of the statistical analysis are shown independently for the indicators of each level of measurement. If not stated in a different way, a complete list with the raw data, mean values and standard deviations is presented in Appendix A.2.

4.3.1. Physiological Assessment

The analysis of physiological measures was based on the same 3-minute-intervals which were used for the DD manipulation. The indicators were submitted to ANOVA for repeated measurements with repetitiveness and sequence of DD as between-subject factors and run and intervals during run (i=16) as within-subject factors. Table 17 gives an overview of the statistical analysis for all indicators (descriptive statistics are presented in Table A-6 to A-9).

Table 17: Results of Analysis of Variance for physiological measures (Study I)

SOURCE	RESULTS ($F_{\text{df hypothesis, df error}}$, p-value)			
	HR (baseline corr.)	**HRV**	**SCL (corr.)**	**No of. Blinks**
REP	***$F_{1,3}$=5.59, p=.099(*)***	$F_{1,3}$=.06, p=.829	***$F_{1,3}$=1.56, p=.174(*)***	$F_{1,3}$=.14, p=.733
DD	$F_{3,3}$ =2.25, p=.262	**$F_{3,3}$=9.42, p=.049***	$F_{3,3}$=.87, p=.327	$F_{3,3}$=.56, p=.678
RUN	**$F_{1,3}$=11.21, p=.044*[a]**	$F_{1,3}$=.92, p=.408	***$F_{1,3}$=9.08, p=.057(*)***	***$F_{1,3}$=4.82, p=.116(*)***
INTER	**$F_{15,45}$=2.82, p=.004**[g]**	$F_{15,45}$=1.24, p=.280[f]	**$F_{15,45}$=9.42, p=.000*** [a]**	**$F_{15,45}$=2.57, p=.007**[g]**
Run x Rep	$F_{1,3}$=.01, p=.930	**$F_{1,3}$=21.02, p=.019* [a]**	***$F_{1,3}$=3.22, p=.171(*)***	$F_{1,3}$=.09, p=.785
Run x DD	$F_{3,3}$=.39, p=.772	***$F_{3,3}$=3.91, p=.146(*)***	***$F_{3,3}$=5.06, p=.108(*)***	$F_{3,3}$=1.80, p=.321
Inter x Rep	$F_{1,45}$=.68, p=.790[g]	$F_{1,45}$=1.15, p=.346[f]	$F_{1,45}$=.74, p=.727 [a]	$F_{1,45}$=1.13, p=.356
Inter x DD	**$F_{45,45}$=1.90, p=.017*[g]**	$F_{45,45}$=1.13, p=.342	**$F_{45,45}$=1.80, p=.025* [a]**	***$F_{4,45}$=1.33, p=.173(*)[g]***
Run x Inter	**$F_{15,45}$=2.95, p=.003**[g]**	$F_{15,45}$=.74, p=.732[g]	$F_{15,45}$=.99, p=.477[c]	***$F_{15,45}$=1.53, p=.136(*)[d]***

[7] Descriptive Data Analysis does not deploy one-sided or two-sided tests; thus, based on the assumption of one-sided testing a p<.10 would result in p<.20 if tested against a two-sided distribution in case a t-test is conducted.

SOURCE	RESULTS ($F_{df\ hypothesis,\ df\ error}$, p-value)			
	HR (baseline corr.)	**HRV**	**SCL (corr.)**	**No of. Blinks**
Run x Inter x Rep	$F_{15,45}$=.70, p=.776^{e}	$F_{15,45}$=.88, p=.594^{g}	$F_{15,45}$=1.28, p=.253^{b}	***$F_{15,45}$=1.38, p=.198(*)^{e}***
Run x Inter x DD	**$F_{45,45}$=2.03, p=.01**^{g}**	$F_{45,45}$=.49, p=.99	$F_{45,45}$=.92, p=.612	$F_{45,45}$=1.21, p=.264^{e}

Note. N=24. Rep=Repetitiveness; DD=Sequence of DD; Inter=Interval during run; HR=heart rate; HRV=heart rate variability; SCL=skin conductance level. Trend effects: ^{a}linear; ^{b}cubic; ^{c}quartic; ^{d}order 5; ^{e}order 7; ^{f}order 8; ^{g}order >8. ***p<.001. **p<.01. *p<.05. (*)p<0.2.

The development of mean HR over time is depicted in Figure 4. The statistical analysis showed that baseline-corrected HR was tendencially lower for the repetitive group compared to the non repetitive group. Mean HR linearly decreased from the first to the second run and fluctuated during the runs, as expressed in the significant higher order trend component. The significant Run x Interval interaction indicated that this decline was more pronounced at the beginning of the first run.

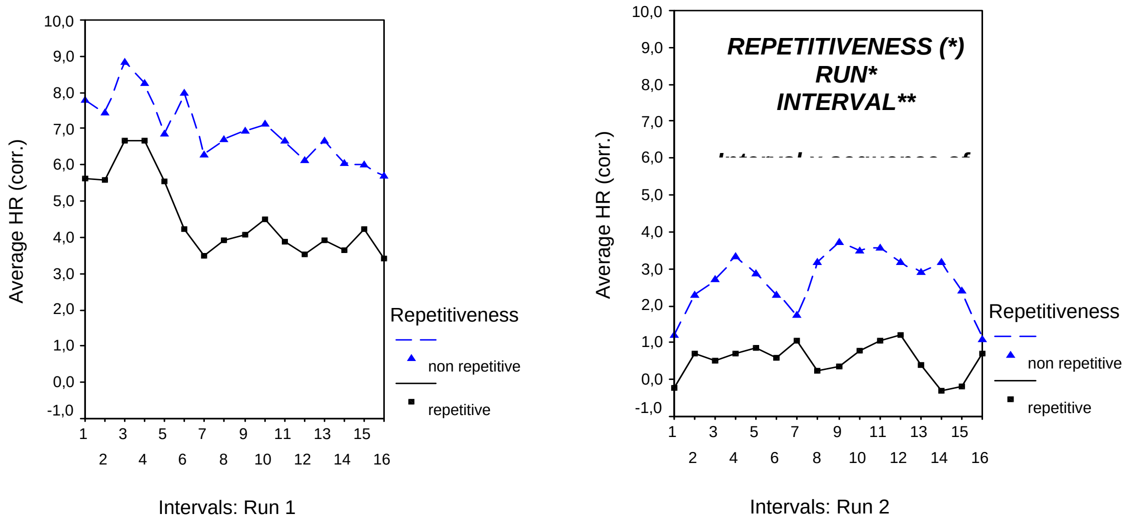


Figure 4: Average heart rate in bpm (baseline-corr.) in 3-minute-intervals for each run as a function of repetitiveness[8].

The significant Run x Interval x DD interaction in Figure 5 reflects a significant higher order trend component and is a consequence of the higher HR under low DD in both runs (l-l), which started towards the end of the first run and remained elevated. No additional significant main or interaction effects were found.

[8] Figures report significant effects according to marks introduced in 4.2.6. In some figures during the document symbols needed to be switched due to software restrictions.

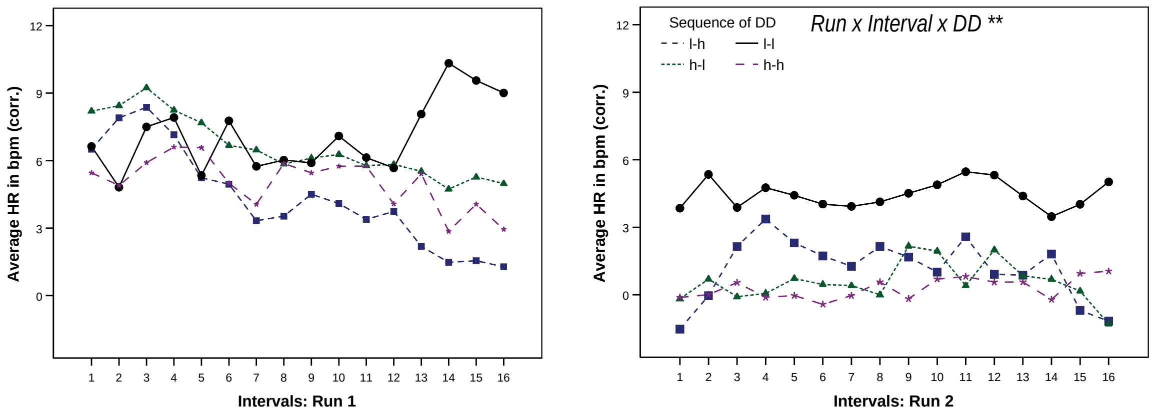


Figure 5: Average heart rate in bpm (baseline-corr.) in 3-minute-intervals for each run as a function of the sequence of Dynamic Density (DD) (Legend: ■ l-h: low DD in Run 1, high density in Run 2. ▲h-l: high DD in Run 1, low DD in Run 2, l-l: •low DD in Run 1 and 2, ✱h-h: high DD in Run 1 and 2).

A significant main effect of the sequence of DD was found in HRV. After a comparison of the descriptive statistics in Table A-4, this effect was mainly attributed to the strong influence of the l-l sequence. The interaction between run and repetitiveness (Figure 6) reflects an increase of HRV in the non repetitive condition of the second run. The marginally significant Run x DD interaction reflected a different development for the two runs, did however not show a different development in any of the trend components. No additional effects were found.

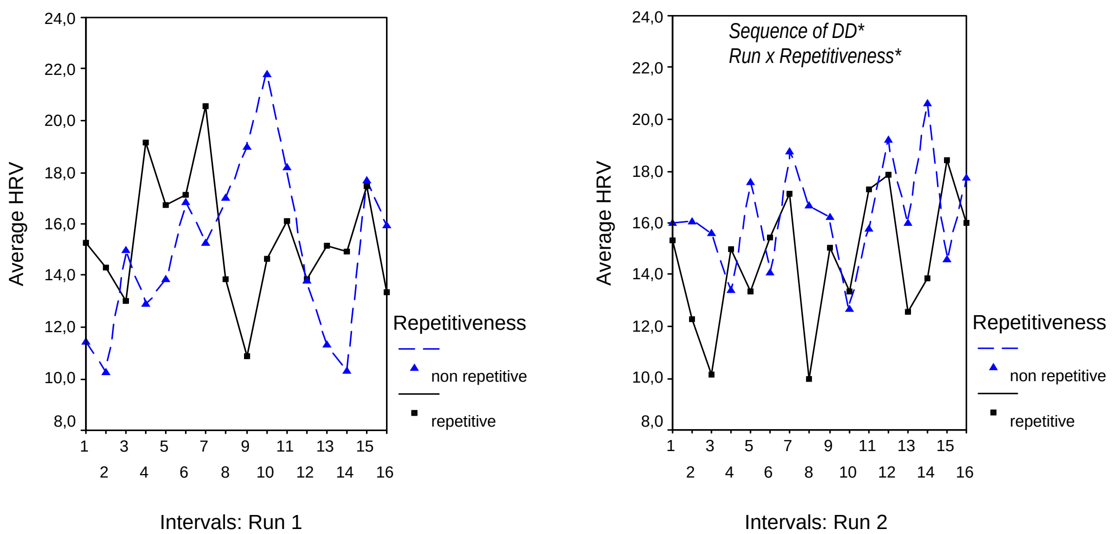


Figure 6: Average heart rate variability in bpm in 3-minute-intervals for each run as a function of repetitiveness

The analysis of SCL resulted in marginally lower values in the repetitive condition, which linearly decreased from the first to the second run. The tendencially significant Run x Repetitiveness interaction reflected the different time course of SCL, which was more pronounced for the repetitive condition especially in the second run (Figure 7). The significant interaction between interval and sequence of DD in SCL indicated again a greater impact of the l-l condition, as Figure 8 reflects.

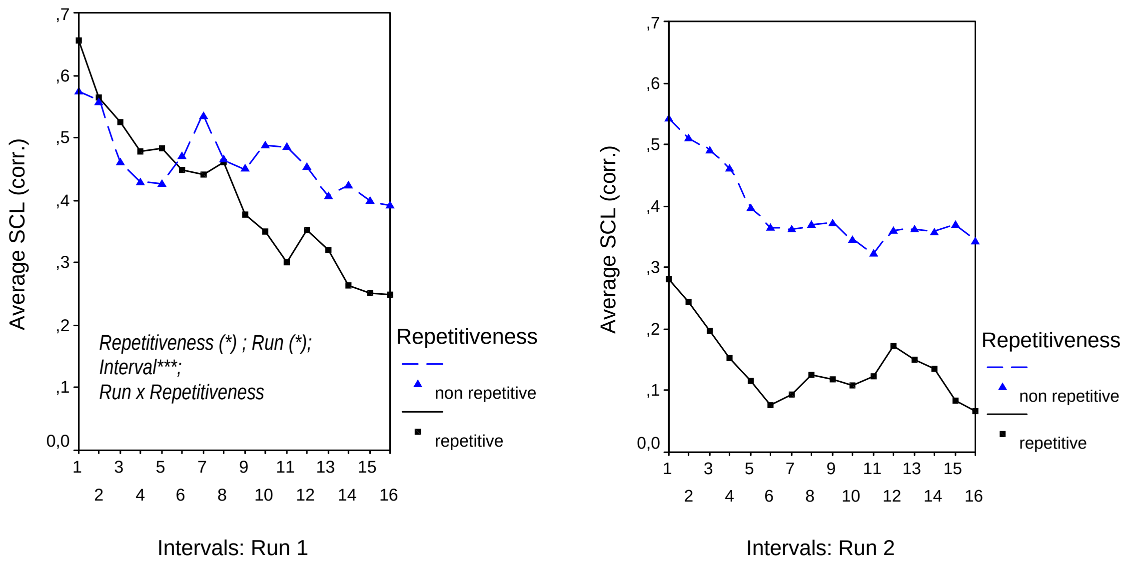


Figure 7: Average skin conductance level in µS (corr.) in 3-minute-intervals for each run as a function of repetitiveness

The increase of SCL occurred later, if compared to the development of the HR in both low density runs.

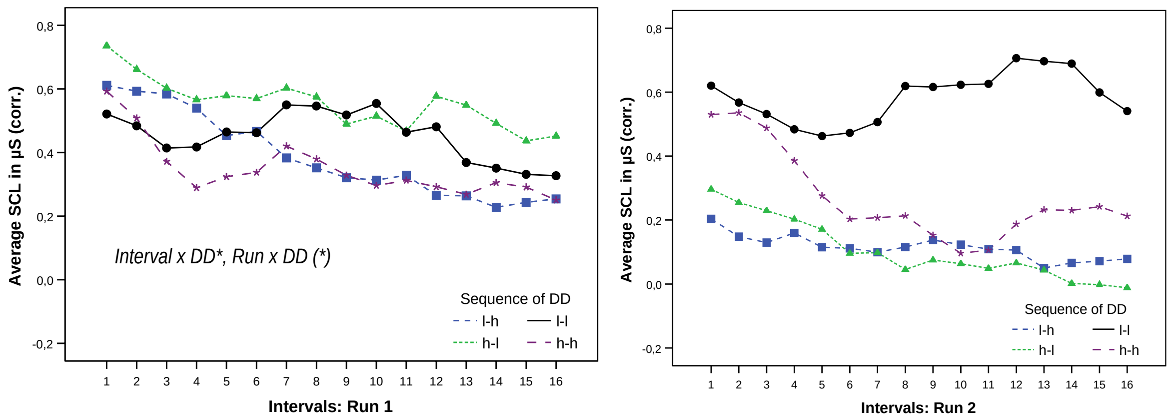


Figure 8: Average skin conductance level (in µS corr.) in 3-minute-intervals for each run as a function of the sequence of Dynamic Density (DD) (Legend: ■ l-h: low DD in Run 1, high density in Run 2. ▲h-l: high DD in Run 1, low DD in Run 2, l-l: •low DD in Run 1 and 2, ✱h-h: high DD in Run 1 and 2)

The marginally significant effects of interval and run on the number of blinks and their interactions indicated differences in the course, which are shown in Figure 9.

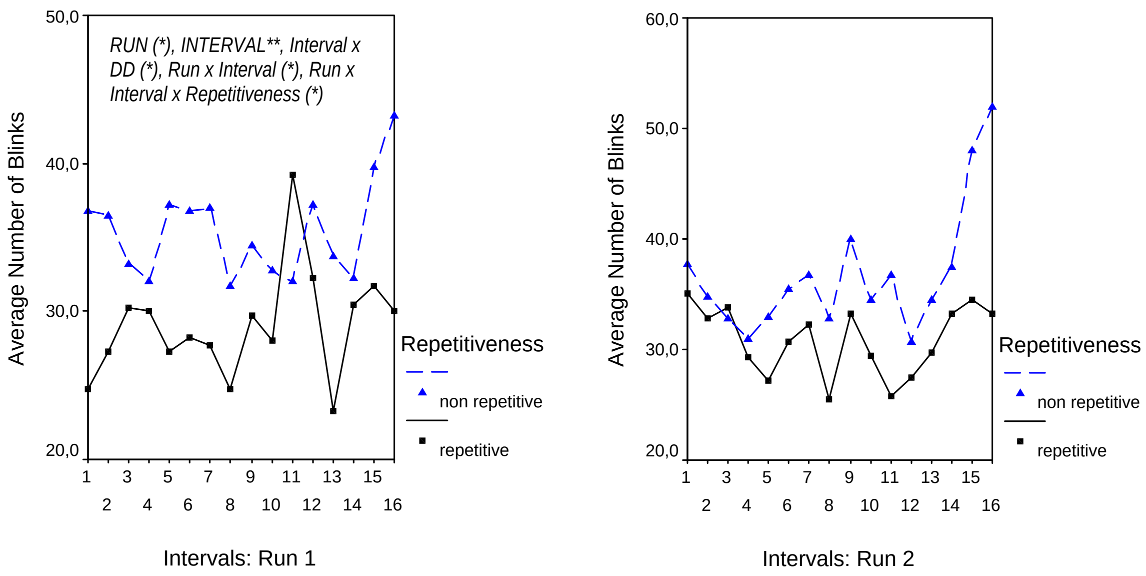


Figure 9: Average number of blinks in 3-minute-intervals for each run as a function of repetitiveness

4.3.2. Subjective Assessment

4.3.2.1. Subjective ratings during the scenario execution (TSI)

A scale to assess motivational, cognitive and emotional aspects was administered at the beginning of each run and at 3 measurement points during the runs. Level-corrected items were submitted to statistical analysis, which included the between-subject factors *repetitiveness* and *sequence of DD* and the within-subject-factors *run* and *interval during run* (i=3). As there was a different interval between the first two points and the additional points of measurement, the one before the run was excluded from analysis because of the interest in computing trends which requires equal intervals. Table 18 contains the results of all statistical analysis; descriptive statistics are presented in Table A-10.

Table 18: Results of Analysis of Variance for subjective ratings during scenarios (Study I)

SOURCE	RESULTS ($F_{\text{df hypothesis, df error}}$, p-value)							
	Attentiveness	Fatigue	Boredom	Irritation	Strain	Concentration	Motivation	Sleepiness
REP	$F_{1,3}$=.49 p=.534	$F_{1,3}$=*4.36* p=*.128(*)*	$F_{1,3}$=.603 p=.494	*$F_{1,3}$=9.39* *p=.055* (*)	*$F_{1,3}$=3.09* *p=.177* (*)	$F_{1,3}$=.03 p= .877	$F_{1,3}$=.41 p=.566	$F_{1,3}$=13.71 p=.034*
DD	$F_{3,3}$=.35 p=.792	$F_{3,3}$=2.25 p=.261	$F_{3,3}$=.146 p=.926	$F_{3,3}$=1.78 p= .323	$F_{3,3}$=.73 p= .601	$F_{3,3}$=.81 p= .567	$F_{3,3}$=.50 p=.710	$F_{1,3}$=25.29 p=.012*
RUN	*$F_{1,3}$=6.82* *p=.08* (*)	$F_{1,3}$=.32 p=.613	$F_{1,3}$=2.27 p=.229	$F_{1,3}$=.16 p=.718	$F_{1,3}$=3.42 p=.162 (*)	$F_{1,3}$=.57 p=.505	$F_{,31}$=1.45 p=.314	$F_{1,3}$=3.27 p=.168 (*)
INTER	*$F_{2,6}$= 3.71* *p=.09* (*)	$F_{2,6}$= 1.45 p=.307	$F_{2,6}$= 1.71 p=.258	*$F_{2,6}$= 4.40* p=*.067* (*)	$F_{2,6}$=1.37 p=.323	$F_{2,6}$=1.82 p=.241	$F_{2,6}$=2.59 p=.155 (*)	$F_{2,6}$=7.64 p=.022*[a]
Run x Rep	$F_{1,3}$=2.45 p=.215	$F_{1,3}$=.00 p=1.00	$F_{1,3}$=.03 p=.878	$F_{1,3}$=.02 p=.903	*$F_{1,3}$=8.44* *p=.062* (*)	$F_{1,3}$=.14 p=.731	$F_{1,3}$=.09 p=.783	$F_{1,3}$=.95 p=.402
Run x DD	$F_{3,3}$=2.45 p=.240	$F_{3,3}$=1.53 p=.360	$F_{3,3}$=2.20 p=.267	$F_{3,3}$=.39 p=.769	F3,3=*3.05* p=*.192* (*)	$F_{3,3}$=.29 p=.835	$F_{3,3}$=.30 p=.824	$F_{3,3}$=1.46 p=.381
Inter x Rep	$F_{2,6}$=.06 p=.943	$F_{2,6}$=.64 p=.56	$F_{2,6}$=.1.29 p=.343	*$F_{2,6}$=3.977* *p=.08* (*)	$F_{2,6}$=.66 p=.550	$F_{2,6}$=.06 p=.943	$F_{2,6}$=.31 p=.744	$F_{2,6}$=.27 p=.770
Inter x DD	$F_{6,6}$=1.35 p=.361	$F_{6,6}$=1.45 p=.330	$F_{6,6}$=6.29 p=.021[b]	$F_{6,6}$=.86 p=.570	$F_{6,6}$=1.98 p=.213	$F_{6,6}$=1.20 p=.417	$F_{6,6}$=.52 p=.779	$F_{6,6}$=1.09 p=.459

SOURCE	RESULTS ($F_{df\ hypothesis,\ df\ error}$, p-value)							
	Attentiveness	Fatigue	Boredom	Irritation	Strain	Concentration	Motivation	Sleepiness
Run x Inter	$F_{2,6}$=.18 p=.842	$F_{2,6}$=.55 p=.604	$F_{2,6}$=.66 p=.551	$F_{2,6}$=1.46 p=.304	$F_{2,6}$=.63 p=.566	$F_{2,6,}$=.07 p=.936	$F_{2,6}$=.78 p=.501	$F_{2,6}$=18 p=.842
Run x Inter x Rep	$F_{2,6}$=.53 p=.614	$F_{2,6}$=.89 p=.460	$F_{2,6}$=.51 p=.623	$F_{2,6}$=2.03 p=.213	$F_{2,6}$=.43 p=.670	$F_{2,6}$=.47 p=.648	$F_{2,6}$=1.44 p=.308	$F_{2,6}$=.18 p=.842
Run x Inter x DD	*$F_{6,6}$=3.00 p=.104* (*)	$F_{6,6}$=.44 p=.830	$F_{6,6}$=.32 p=.902	$F_{6,6}$=.37 p=.876	$F_{6,6}$=.85 p=.578	$F_{6,6}$=1.67 p=.275	$F_{6,6}$=.93 p=.536	$F_{6,6}$=.29 p=.919

Note. N=24. Rep=Repetitiveness; DD=Sequence of Dynamic Density (DD); Inter=Interval during run. Trend effects: [a]linear; [b]quadratic.
***p<.001. **p<.01. *p<.05. (*)p<0.2.

The main effect of repetitiveness revealed higher sleepiness in the repetitive condition, as depicted in Figure 10. As the (marginally) significant effects indicate, sleepiness increased from the first to the second run and linearly during the scenarios. The main effect of sequence of DD is a consequence of increased ratings in the l-l condition (Figure 11).

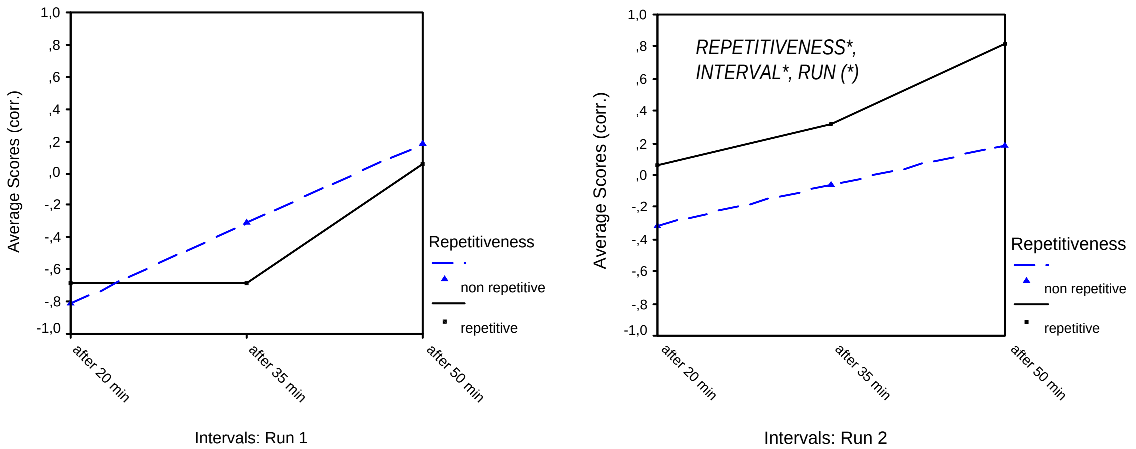


Figure 10: Average ratings of sleepiness (level-corr.) for each run as a function of repetitiveness

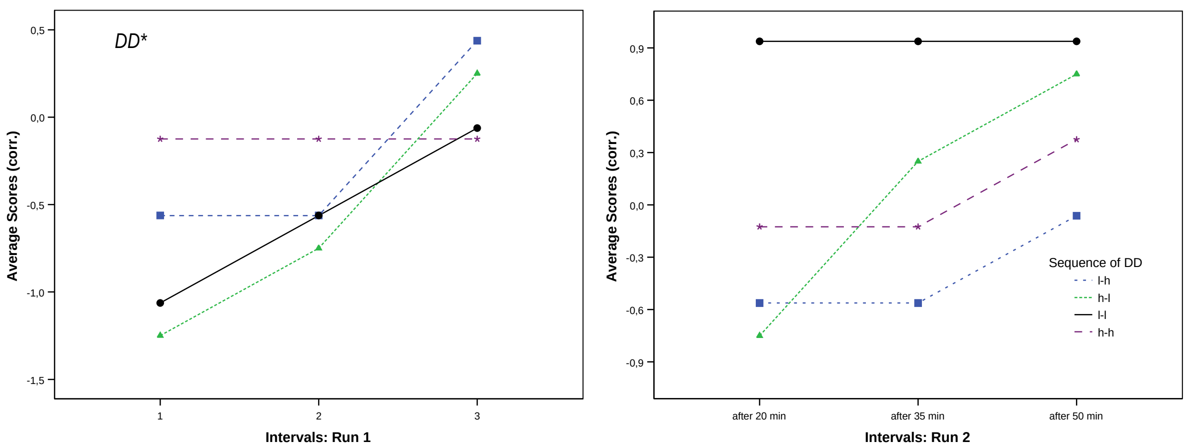


Figure 11: Average ratings of sleepiness (level-corr.) for each run as a function of the sequence of Dynamic Density (DD) (Legend: ■ l-h: low DD in Run 1, high density in Run 2. ▲h-l: high DD in Run 1, low DD in Run 2, l-l: •low DD in Run 1 and 2, ✱h-h: high DD in Run 1 and 2)

For fatigue-scores, marginally significant differences were found between treatment conditions. Participants felt more fatigued in the repetitive condition, as can be seen in Figure 12. A comparison with the sleepiness items reveals a similar development.

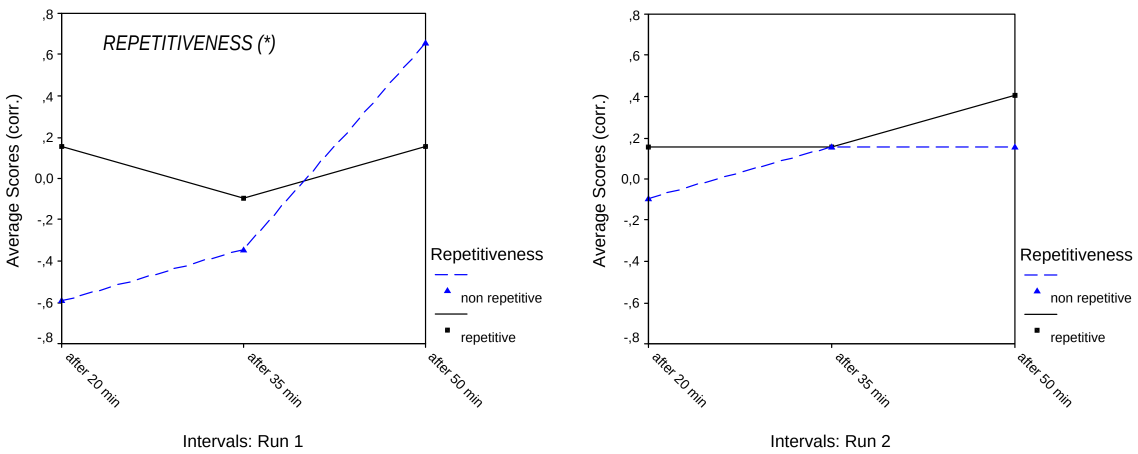


Figure 12: Average ratings of fatigue (level-corr.) for each run as a function of repetitiveness

For attentiveness, a tendencially significant interaction between sequence of DD, run and interval resulted. Attentiveness decreased from the first to the second run and within the scenarios. At the beginning of a scenario the repetitive group felt less attentive, while at the end and after a pronounced decrease the non repetitive group rated lower attentiveness (Figure 13).

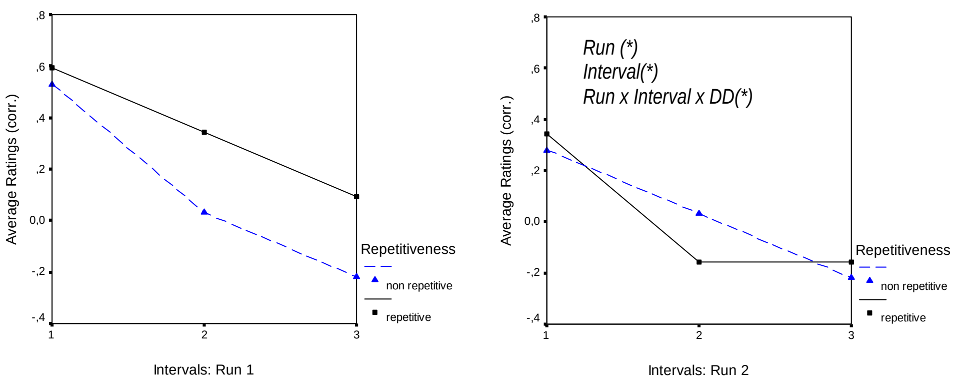


Figure 13: Average ratings of attentiveness (level-corr.) for each run as a function of repetitiveness

Strain was tendencially rated lower in the repetitive condition; however, the Run x Repetitiveness interaction presented in Figure 14 reveals a different course through higher strain of the non repetitive group in the first, but of the repetitive group in the second run. The interaction between run and DD was marginally significant.

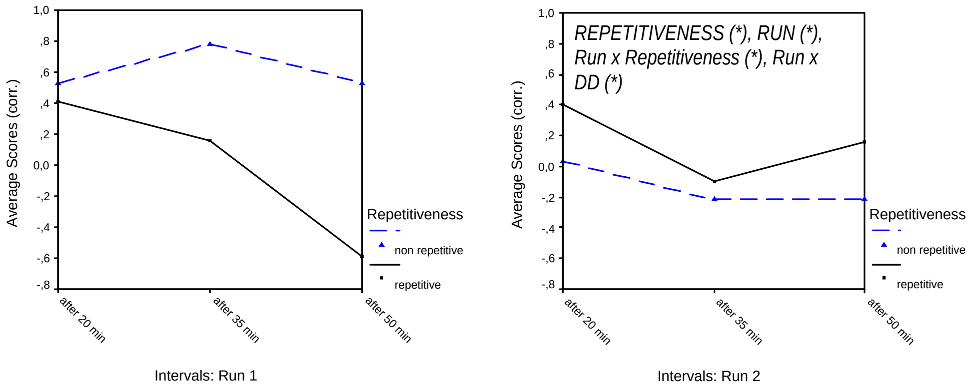


Figure 14: Average ratings of strain (level-corr.) for each run as a function of repetitiveness

Participants felt tendencially more irritated in the non repetitive condition. This was however different at the beginning of the first run, where the repetitive group felt higher irritation. In the boredom-ratings (Figure 15) there was an interaction between interval and sequence of DD that followed a quadratic trend. This was an effect of the higher boredom in the h-h condition. No additional main or interaction effects were found in boredom, or in motivation and concentration.

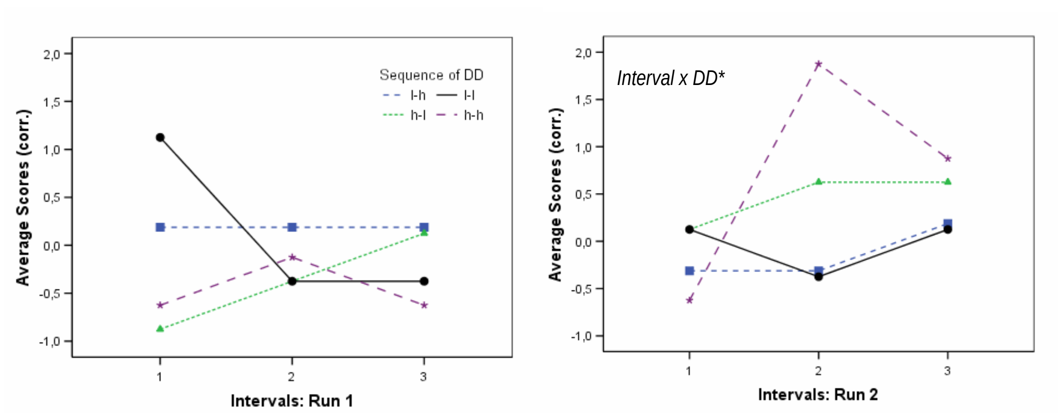


Figure 15: Average ratings of boredom (level-corr.) for each run as a function of the sequence of Dynamic Density (DD) (Legend: ■ l-h: low DD in Run 1, high density in Run 2. ▲h-l: high DD in Run 1, low DD in Run 2; l-l: •low DD Run 1+2, ✱h-h: high DD Run 1 + 2).

4.3.2.2. Workload assessment after scenario (NASA-TLX)

Table 19: Results of Analysis of Variance for workload measures (Study I)

SOURCE	**RESULTS ($F_{df\ hypothesis,\ df\ error}$, p-value)**							
	Mental demand	Physical demand	Temporal demand	Effort	Performance	Frustration	Feeling of Monotony	Overall
REP	$F_{1,3}$=1,96, p=.256	$F_{1,3}$=1.04, p=.383	***$F_{1,3}$=3.37, p=.164*** (*)	$F_{1,3}$=2.23, p=.232	$F_{1,3}$=1.28, p=.340	$F_{1,3}$=1.68, p=.286	**$F_{1,3}$=12.34 p=.039***	***$F_{1,3}$=3.01, p=.181*** (*)
DD	$F_{3,3}$=.37, p=.783	$F_{3,3}$=2.26, p=.260	$F_{3,3}$=.70, p=.611	$F_{3,3}$=.77, p=.582	$F_{3,3}$=.23, p=.871	$F_{3,3}$=1.68, p=.920	$F_{3,3}$=.15, p=.926	$F_{3,3}$=.15, p=.926
RUN	$F_{1,3}$=.531, p=.50	$F_{1,3}$=.17, p=.709	$F_{1,3}$=.14, p=.738	$F_{1,3}$=.21, p=.677	$F_{1,3}$=.62, p=.487	$F_{1,3}$=2.41, p=.218	$F_{1,3}$=.83, p=.431	$F_{1,3}$=.03, p=.867
Run x Rep	$F_{1,3}$=1.78., p=.274	$F_{1,3}$=.00, p=.961	$F_{1,3}$=.26, p=.646	***$F_{1,3}$=6.06, p=.091*** (*)	$F_{1,3}$=.21, p=.679	$F_{1,3}$=.36, p=.592	$F_{1,3}$=1.93, p=.259	$F_{1,3}$=.54, p=.516

SOURCE	RESULTS ($F_{df\ hypothesis,\ df\ error}$, p-value)							
	Mental demand	Physical demand	Temporal demand	Effort	Performance	Frustration	Feeling of Monotony	Overall
Run x DD	$F_{3,3}$=.68, p=.621	$F_{3,3}$=1.17, p=.449	$F_{3,3}$=1.58, p=.359	***$F_{3,3}$=5.97, p=.088*** (*)	$F_{3,3}$=1.35, p=.405	$F_{3,3}$=2.35, p=.251	$F_{3,3}$=1.67, p=.341	$F_{3,3}$=1.50, p=.374

Note. Rep=Repetitiveness; DD=Sequence of Dynamic Density; ***p<.001. **p<.01. *p<.05. (*)p<0.2.

Table Table 19 contains the results of the ANOVA for the NASA-TLX items including a summarized workload indicator and the ratings for the feeling of monotony. Descriptive statistics are presented in Table A-11.

The marginally significant main effects of repetitiveness in temporal demand and overall workload revealed higher ratings in the non repetitive group, while the subjective feeling of monotony was rated significantly higher in the repetitive group (Figure 16).

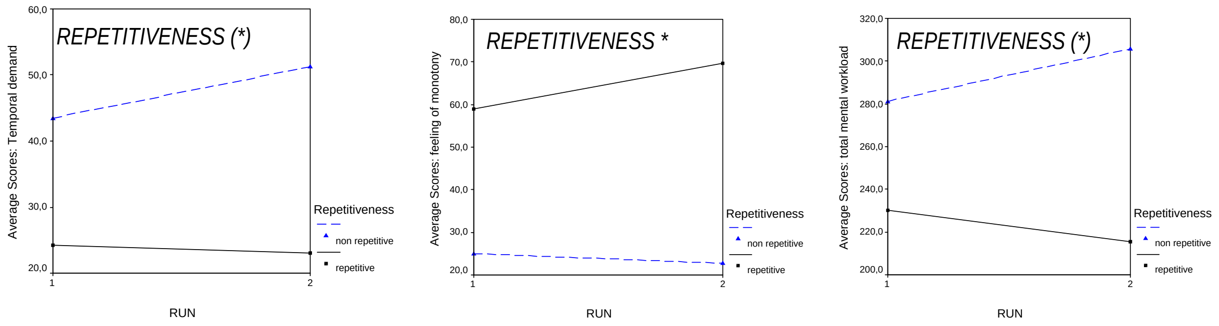


Figure 16: Average ratings of temporal demand, feeling of monotony and total workload for each run as a function of repetitiveness

The significant interaction between run and repetitiveness and run and sequence of DD for the perceived effort is demonstrated in Figure 17. Increased effort was perceived in the second run after low DD in the first run that changed to high or remained low.

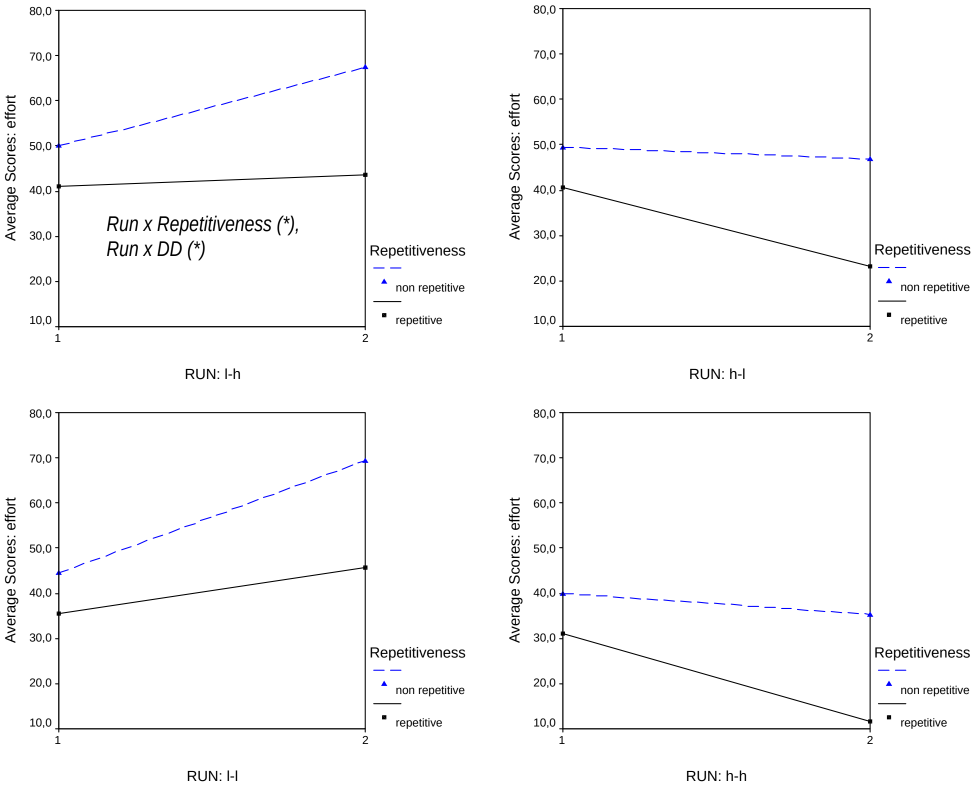


Figure 17: Average ratings of effort for each run as a function of repetitiveness and sequence of dynamic density (DD) (Legend: l-h: low DD in Run 1, high density in Run 2. h-l: high DD in Run 1, low DD in Run 2, l-l: low DD in Run 1 and 2, h-h: high DD in Run 1 and 2.)

4.3.2.3. Mood assessment after scenario (UWIST Mood Assessment scale)

Table 20 represents the results of the statistical analysis for the mood subscales (descriptive statistics in Table 12).

Table 20: Results of Analysis of Variance for assessment of critical states (Study I)

SOURCE	RESULTS ($F_{df\ hypothesis,\ df\ error}$, p-value)		
	Hedonic Tone	**Energetic Arousal**	**Tense Arousal**
Rep	$F_{1,3}$=.08, p=.793	$F_{1,3}$=.68, p=.469	***$F_{1,3}$=4.06, p=.137*** (*)
DD	$F_{3,3}$=.08, p=.967	$F_{3,3}$=.29, p=.831	$F_{3,3}$=1,37, p=.401
Run	$F_{1,3}$=.00, p=1.00	$F_{1,3}$=2.46 p=.215	$F_{1,3}$=.04, p=.861
Run x Rep	***$F_{1,3}$=3.74, p=.149*** (*)	***$F_{1,3}$=5.12, p=.109*** (*)	***$F_{1,3}$=2.93, p=.186*** (*)
Run x DD	$F_{3,3}$=.47, p=.722	$F_{3,3}$=1.57, p=.361	$F_{3,3}$=1.20, p=.444

Note. Rep=Repetitiveness; DD=Sequence of DD; Inter=Interval. ***p<.001. **p<.01. *p<.05. (*)p<0.2.

Tense arousal was tendencially higher in the non repetitive condition, but in the second run the marginally significant interaction with run indicated that they were approaching each other.

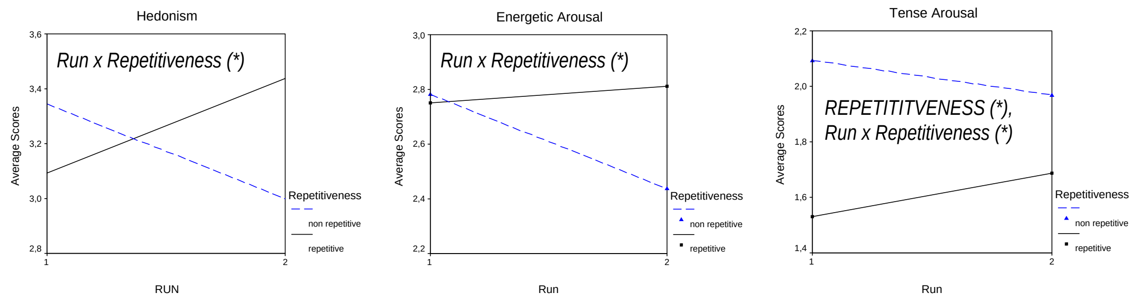


Figure 18: Average scores in UWIST mood assessment subscales for each run as a function of repetitiveness

The additional interaction between run and repetitiveness reflected a decline in energetic arousal for the non repetitive condition. The interaction between run and repetitiveness in hedonic tone showed that the repetitive condition was perceived more pleasant in the second run.

4.3.2.4. The assessment of critical states after scenario (SOF)

The results for the SOF subscales are shown in Table 21. Information on descriptive statistics is contained in Table 13. The mean stress scores demonstrate that tendencially more stress was perceived in the non repetitive condition. Satiation increased significantly and monotony increased marginally from the first to the second. No effect was found in the fatigue subscale.

Table 21: Results of Analysis of Variance for assessment of critical states (Study I)

SOURCE	RESULTS ($F_{df\ hypothesis,\ df\ error}$, p-value)			
	Stress	**Monotony**	**Fatigue**	**Satiation**
REP	***$F_{1,3}$=3.64, p=.152*** (*)	$F_{1,3}$=1.19, p=.354	$F_{1,3}$=.231, p=.664	$F_{1,3}$=.00, p=1.00
DD	$F_{3,3}$=.19, p=.901	$F_{3,3}$=.67, p=.623	$F_{3,3}$=.062, p=.977	$F_{3,3}$=.292, p=.831
RUN	$F_{1,3}$=.15, p=.728	***$F_{1,3}$=4.41, p=.127*** (*)	$F_{1,3}$=2.43, p=.217	**$F_{1,3}$=11.52, p=.043***
Run x Rep	$F_{1,3}$=.17, p=.709	$F_{1,3}$=2.11, p=.242	$F_{1,3}$=.57, p=.507	$F_{1,3}$=2.46, p=.215
Run x DD	$F_{3,3}$=1.29, p=.420	$F_{3,3}$=.79, p=.574	$F_{3,3}$=1.54, p=.365	$F_{3,3}$=2.79, p=.210

Note. Rep=Repetitiveness; DD=Sequence of DD; Inter=Interval. ***p<.001. **p<.01. *p<.05. (*)p<0.2.

The development of the subscale scores is presented in Figure 19. That no main effect was found in the monotony subscale, will be further discussed in the next section.

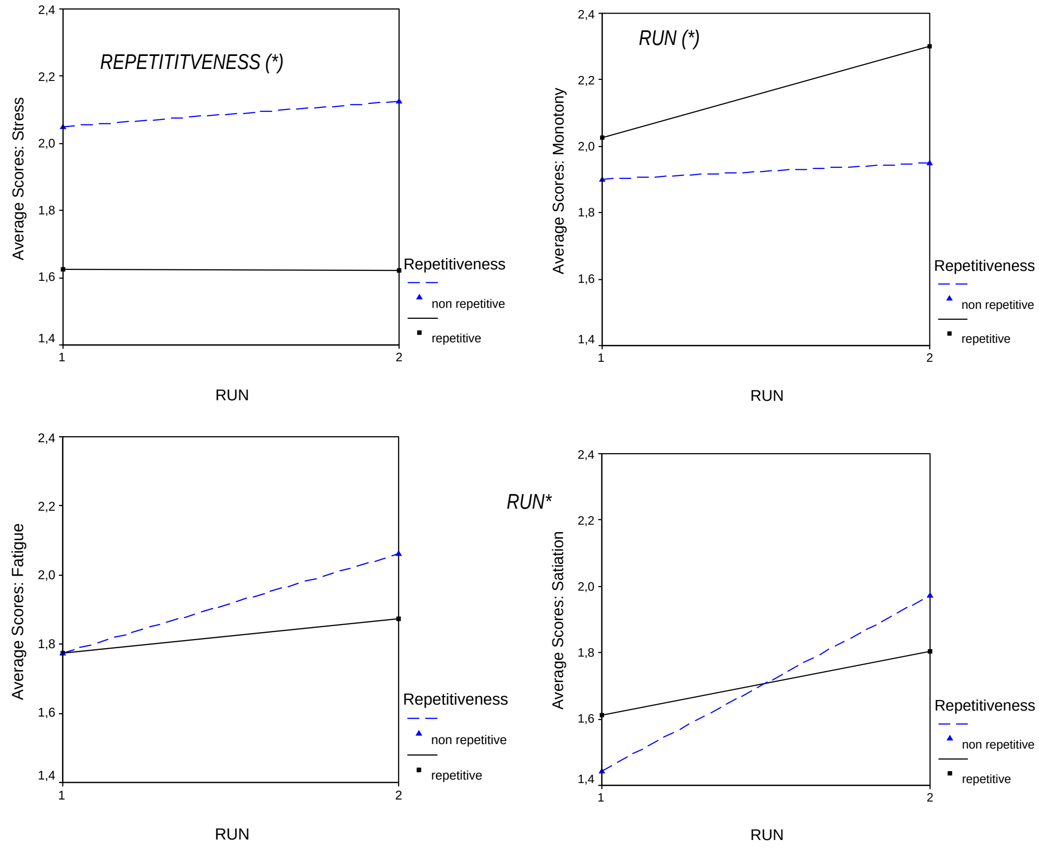


Figure 19: Average scores in Scale of Feelings(SOF) subscales for each run as a function of repetitiveness

4.3.2.5. The assessment of situational awareness after scenario (SASHA)

There were no significant differences or interactions in the ratings of the SASHA-items, marginally significant effects are presented with the descriptive indicators in Tables A-14 and A-15 and indicate that controllers were less sure if they forgot to transfer aircraft (item 5) in the non repetitive condition. The marginally significant interaction between run and sequence of DD reflected the impact of the I-I condition with a markedly decreased rating for the difficulty of finding information (item 6) in the second run.

4.3.3. Performance Assessment

The analysis of STCA alerts in each scenario revealed two STCA alerts in the first run in the non repetitive setting, one of low, one of high DD. In the second run, in non repetitive condition two STCA alerts occurred in low density and one in high DD, one STCA alert occurred in the repetitive scenario of low traffic. Table 22 contains the summarized values which were submitted to Fisher's exact test. There is no significant difference between the repetitiveness (p=.179) and the DD conditions (p=.654).

Table 22: Frequency of STCA events (STCA/ No STCA) for out-of-routine conflict situation (Study I)

		DD		
		Low	High	Total
		STCA/ No STCA	STCA/ No STCA	STCA/ No STCA
Repetitiveness	Repetitive	1/7	0/8	1/15
	Non repetitive	3/5	2/6	5/11
	Total	4/12	2/14	6/26

Note. (N=8, 2 Scenarios)

The Vienna Test System was used to assess performance after the traffic scenarios were completed. There were no significant main effects of repetitiveness or sequence of DD on median reaction time or motor time. There was tendencially more variation in motor time under the repetitive condition (Table 23 descriptive statistics Table A-16). Due to data loss, the results of the Cognitrone were not considered; the raw data is presented in Tables A-17.

Table 23: Results of Analysis of Variance for performance assessment (Study I)

SOURCE	RESULTS (F $F_{df\ hypothesis,\ df\ error}$, p-value)			
	RT Reaction time	Motor Time	Distribution reaction time	Distribution motor time
REP	$F_{1,2}$=0.20; p=.685	$F_{1,2}$=0.54; p=.516	$F_{1,2}$=0.69; p=.467	***$F_{1,2}$=3.41; p=162$^{(*)}$***
DD	$F_{3,2}$=0.59; p=.660	$F_{3,2}$=0.50; p=.708	$F_{3,2}$=2.39; p=.246	***$F_{3,2}$=4.21; p=134$^{(*)}$***

***Note*. Rep=Repetitiveness; DD=Sequence of DD. ***p<.001. **p<.01. *p<.05. (*)p<0.2.**

The analysis of the test for time-movement anticipation ZBA (descriptive statistics in Table A-18) revealed no significant effects in the total median deviation time (REPETITITVENESS: $F_{1,2}$=0.61, p=.517; DD: $F_{3,2}$=1.23, p=.478) nor in the direction deviation (REPETITITVENESS: $F_{1,2}$=0.87; p=.449; DD: $F_{3,2}$ =1.00, p=.535).

4.4. DISCUSSION

The aim of the preliminary study was to examine empirically whether earlier identified task factors evoke a state of monotony with the current set-up. It was hypothesized that differences in subjective, physiological, and behavioral measures depend on repetitiveness and sequence of dynamic traffic density. The findings suggest that these characteristics affect several physiological and subjective measures; however, no effects on after-task performance could be determined with the used indicators. The following subsections discuss the results in detail. In addition, the procedure for the planned main experiment was revised.

4.4.1. Physiological, Subjective and Behavioral Effects

The results provide support for physiological deactivation as a consequence of repetitiveness. On the physiological level, the tendencially significant effect of repetitiveness indicated that heart rate (HR) was lower in repetitive traffic. In addition, HR reduced from the first to the second run and also developed differently during each run. Significant interactions with the sequence of dynamic density (DD) revealed a different course where especially the condition of low density in both runs indicated increased HR towards the end of the first run and higher values during the second run.

The significant effect of DD on heart rate variability (HRV) may be a consequence of the strong impact of the l-l sequence in the first and second run. The interaction between run and repetitiveness in HRV (Figure 20) reflects a different development depending on repetitiveness, resulting in lower HRV in the repetitive condition, also moderated by the influence of the l-l condition. The increase of HRV in the non repetitive condition from the first to the second run might be explained by getting used to the task. An alternative interpretation attributes this effect to fatigue. A similar pattern in HR and SCL support this assumption. That blinks increased towards the end of each run, might also indicate fatigue according to Stern, Boyer, and Schroeder (1994), since blink rate is a well established measure.

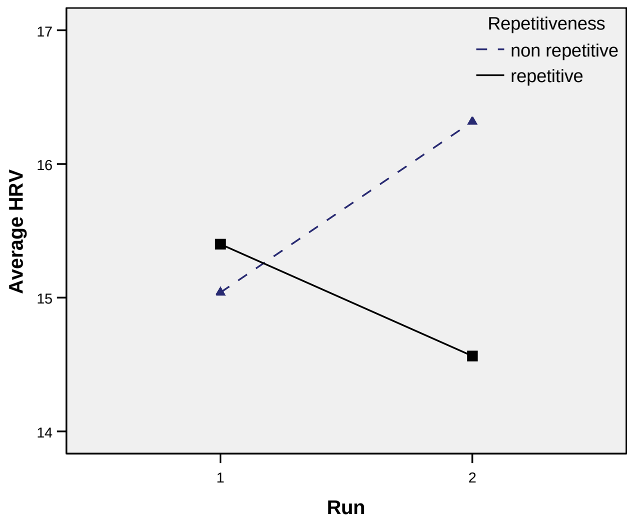


Figure 20: Interaction between run and repetitiveness in HRV (Legend: ■ repetitive, ▲ non repetitive)

The peak values of the HRV in the non repetitive group during the first run cannot be attributed to individual outliers (Figure 21), or to increased task demands that are usually reflected in decreased HRV. On the contrary, outliers in the higher repetitive ratings of the first and last third of the run might have resulted in the higher indicators for repetitive traffic. The increased blink rate at the same time might however indicate that an increased effort to counteract suboptimal activation might have occurred, which was not obvious to the participants. The ratings of sleepiness and strain support this assumption. At the same time, HR slightly increased. Efforts to increase the activation level might have occurred to prevent the performance breakdown. This explanation is plausible, as the occurrence of performance impairments occurs already at an early stage, a result frequently reported in vigilance experiments.

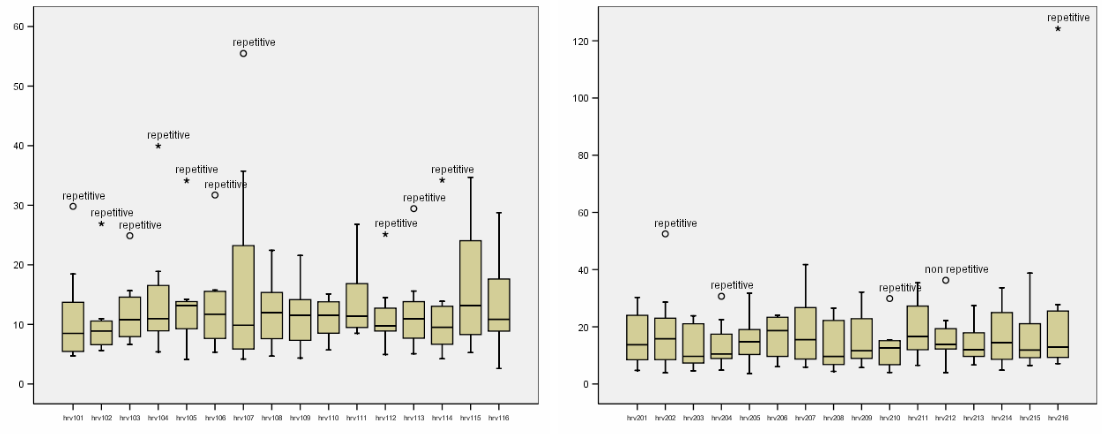


Figure 21: Boxplot of the HRV indicators for the intervals in the first (left graph) and second (right graph) run

The analysis of skin conductance level (SCL) only showed a tendencially significant difference between the repetitive and non repetitive groups because of the initial similarity of the course in the first half of the first run, reflected in the main effect of intervals during run. There was a different development if the first is compared to the second run, where SCL was clearly decreased in the repetitive condition. Similar to the HR, SCL decreased from the first to the second scenario.

There was a similar course of the values in HR and SCL for the condition l-l indicating increasing activation if DD remained low during the second run. At the same time, sleepiness was rated higher in this group. This development is explained to reflect efforts to counteract suboptimal activation, which resulted in increased strain towards the end of the second run. The effects of time-on-task were reflected in significant main effects and interactions and were not unexpected in relation to commonly occurring habituation. Remarkable is as well the deactivation reflected in the SCL decrease from the first to the second run if the DD changed from high to low.

The subjective indicators presented a similar pattern compared to the physiological parameters. Subjective sleepiness and fatigue increased from the first to the second run and attentiveness decreased. A significant effect of repetitiveness was only found in sleepiness, where controllers felt sleepier in repetitive scenarios. However, the ratings of strain show an opposite result as participants felt more strained in the non repetitive and less strained in repetitive traffic. This finding can be seen in relation to the higher number of STCAs that partly occurred because of problems with the simulation environment, as mentioned by the participants. Another explanation is that strain has a negative connotation and might as well be connected to the perceived traffic load. At the same time only marginally significant differences in workload assessed by NASA-TLX were found. Only the subjective feeling of monotony reflected a very clear effect of repetitiveness with higher values under repetitive conditions in both runs, there was however no significant effect in the monotony subscale (compare Table 21).

Not as clear was the effect on self-rated boredom, which supports the distinction between the concepts of monotony and boredom. Moreover, it is noted that no differences in motivation and concentration were found. Apparently, the meaning of attentiveness and concentration was different for the controllers. An interesting finding was also that repetitiveness was affecting irritation but not motivation. At the same time, the scores in the satiation scale were higher in the second run. To reflect these findings within the satiation concept, one would expect that increased irritation would be combined with decreased motivation, which was not the case in the current sample.

Generally, the participants exposed to unvaried low density in both runs showed the most extreme values. This would generally fit in the explanation of underload as stress. In the current case there is no indication, if these two participants were having alternative strategies to increase their activation. Increased effort might have been exerted while fatigue was getting stronger; fatigue was also reflected in the increasing scores from the first to the second run. Also, it might be because of the unexpected event, reflected in the increasing HR in the second run, that the indicators were increased in that run. If the interaction between run and DD is regarded, effort was increasing from the first to the second run in the l-l and l-h condition. This means that after low traffic periods, more effort is perceived if task execution continues.

The higher ratings in the stress subscale for assessing critical states indicate a similar picture compared to the ratings of temporal demand. That the subscale of monotony did not reveal a clear result compared to the ratings in the subjective feeling of monotony might be a consequence of the widespread meanings of the items included in the scale. If further inspected, the items contain too much variety as they refer to the task characteristics which can be distinguished in terms of repetitiveness and uneventfulness. At the same time the items do address feelings and states. For this reason, a review of this scale is necessary in the following research activities. In the mood states only energetical and tense arousal were affected by interactions between repetitiveness and run. Tense arousal was lower and energetic arousal was increased under repetitiveness. An increased energetic arousal seems to reflect the impact of the DD sequence, which was already described for other indicators. But the results might also be explained by the fact that the questionnaire was elicited after the end of the scenarios. But since tense arousal and irritation have a similar pattern, this aspect might be of minor importance.

The task effects on performance measures were not clear to deduct assumptions for a main indicator on monotony. Because of mentioned limitations with the simulation environment, the descriptively higher number of STCA alerts in the non repetitive condition can be explained with these restrictions. At the same time, the higher number of the STCA alerts in the lower density condition can be partly explained by underloading traffic situations. One classical example was reported by one participant, who missed a conflict in the repetitive condition of low traffic density in the second run. In the description of the situation the participant reported to sleep, he was not really tired, but he had the feeling that it was the same all the time. In after-task performance a tendency of increased variation was reflected in the motoric component of the reaction test. However, the null hypothesis concerning the effect of task factors on after-task performance indicators was maintained for all other indicators. Due to missing values, the statistical analysis of concentration indicators was not possible. It is noted that the standard deviations appear to be greater in the repetitive than in the non repetitive condition. If this characteristic is observed again in the main study, it would support the variation in the task performance.

In summary, the results confirm the assumption that repetitive tasks lead to changes as found in monotony studies reported in the literature. Overall, the findings are in line with results presented by Thackray et al. (1975), who observed increased HRV and reduced HR in the monotony/ boredom group as well as increased sleepiness. The results contradicted the reported ratings of strain. Also, no impact was found on HRV, what was a consequence of a different development in the varied dynamic density conditions. Also, the frequently reported reduced attentiveness in vigilance studies was found. To conclude, the introduction of repetitiveness and different levels of DD confirmed the assumption of resulting monotony on different levels as established in the literature and thus can be maintained for the main study.

4.4.2. Methodological Issues

The design of the repetitiveness factor was successful as far as the ratings of the subjective feeling of monotony were concerned. Weaknesses can be discussed for the selected questionnaires and scales. As it turned out, the questionnaire to assess situation awareness did not reveal any relevant differences between conditions. This does not mean, that situation awareness was unaffected by the experimental situation, as can be concluded from individual statements. That fatigue was not reflected in the SOF might also have been an effect of the experimental set-up. Also, this questionnaire is known to not well-reflect critical states within shorter periods. The single items as already successfully applied in the study of Thackray, were more reliable. Also, some of the items of the SOF scale were described as too complicated by participants. Another explanation for the contradictory results reflected in the SOF and single items on fatigue is the fact that the questionnaire was provided after the end of the scenarios, when already some recovery had taken place. The development of the item ratings supports the assumption that the interruption of the task was not perceived as a short rest-break discussed by Schuette (1999), as it was not long enough to break the annoyance of the work. The intervals between the scales and the intervals for the physiological parameters are suitable to reflect the task effects. The course of HR and SCL is more stable and less varying, so that the intervals might have been aggregated, however, the fine grained analysis helped in the interpretation of individual effects. For example, it reflected the increase of HR in the second run after the DD changed from low to high.

When interpreting the multiple reaction patterns it should be taken into account that the sample consisted of operational experts which were not actively working as air traffic controllers any more. This might be relevant concerning the automatic reaction to certain situations as well as habitual procedures to cope with task demands. In addition, the time of day was held constant for the experiment. The afternoon schedule was interpreted to have an influence on the fatigue, as explained in the debriefings. Also, individual impairments at the beginning of the session might have influenced the results, as the state of recovery was lower for the non repetitive group. The traffic scenarios were perceived between easy and moderately heavy depending on the condition. The collection of physiological measures was perceived positively as it was not physically uncomfortable or affecting the task. It was criticized that the breaks between the scenarios were short, which did not reflect usual working days.

Because of the n=1 for each condition, the interaction between the two between-subjects factors was not automatically calculated by the statistical program. Graphical inspection of the interaction between repetitiveness and sequence of DD did not indicate any different developments in any indicators. For this reason, a manual calculation as proposed in Bortz (1999, p. 314 ff.) was not conducted, since this interaction was of low practical importance. No impact of the interaction in the main study was expected, where the analysis of trend effects in DD was considered to be of major importance for the formulation of the experimental hypothesis.

The sample size of n=8 was considered. In that case, the power of the preliminary study is low by definition, even though the application of a mixed design with repeated measurements contributed to increase the power. However, to estimate if the experimental manipulation was appropriate, an n=4 for the most important effect was sufficient to check if there might be differences. Other issues of the simulation which were expected like increased risk seeking where not observed during the experiment.

4.4.3. Consequences for the Main Simulation Study

4.4.3.1. The selection of an indicator for the state of monotony

The results of this investigation provide some initial empirical justification for the effects of task characteristics. They were found on multiple measurement levels which underlines the importance of their consideration. For this reason, the strongest indicators from each level were combined to an indicator for the state of monotony. The goal of this prestudy was to find the most appropriate indicators for a state of monotony to formulate hypotheses for a main study. As such, it is possible to decide which of the indicators can be used to formulate confirmative hypotheses for the main study. Concerning the impact of repetitiveness there are differences depending on the traffic density. The strongest effects were reflected in HR, sleepiness ratings and the subjective feeling of monotony. Therefore, these indicators were z-standardized and combined in a composed indicator for the state of monotony, as supported from a theoretical perspective (Formula 1 in Appendix A). This composition of individual indicators to a main indicator also has the advantage to reduce the number of statistical analysis conducted and meets the proposition of Rosnow and Rosenthal (1989). The consideration of such a composed indicator in the present study for the statistical analysis reveals a significant main effect on repetitiveness ($F_{1,3}$=.00, p=.019, η_p^2=.80), however, no significant interaction between run and sequence of DD (Figure 22), which is again a result of the influence of the low DD sequence. Hence, it is concluded that this indicator can be applied in confirmative hypothesis tests.

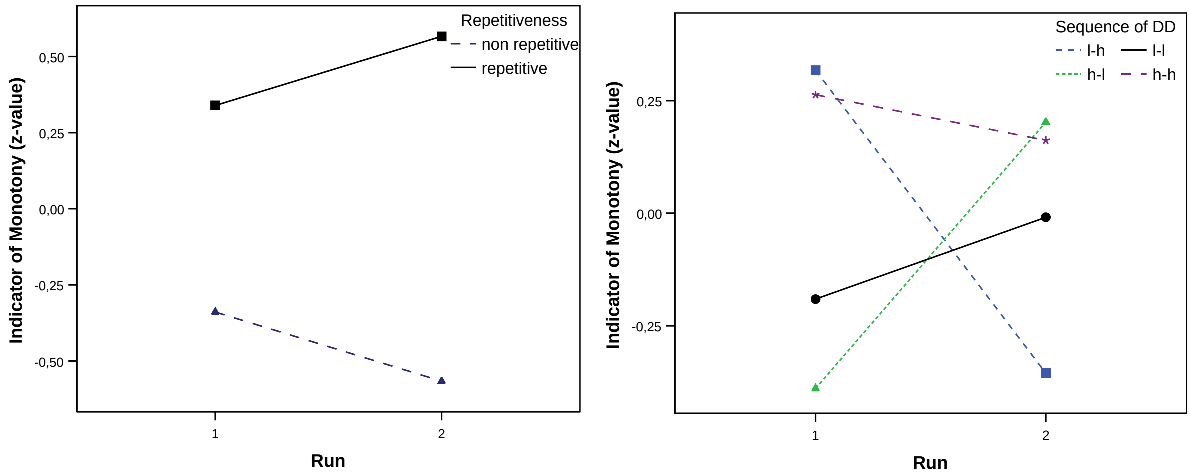


Figure 22: The average values of a composed indicator for the state of monotony as a function of repetitiveness in the left graph (Legend: ■ repetitive; ▲non repetitive) and of the sequence of Dynamic Density (DD) in the right graph (Legend: ■ l-h: low DD in Run 1, high density in Run 2. ▲h-l: high DD in Run 1, low DD in Run 2, l-l: •low DD in Run 1 and 2, ✱h-h: high DD in Run 1 and 2)

It can be argued that the indicator for the state of monotony should be uni-dimensional and thus meet the criteria of high intercorrelations. The empirical data does not support this assumption. Internal consistencies of the indicators reveal a Cronbach's Alpha of r=.36 for the first run and of r=.52 for the second run. Considering the corrected item-total correlation, a different contribution of the items for the two measurement points is noted (Run 1: HR r=.34, monotony r=.10, sleepiness r=.12; Run 2: HR r=.15, monotony r=.04, sleepiness r=.12). This might reflect a different development of the indicators depending on time-on-task. For this reason, the multivariate analysis of variance (MANOVA) was preferred for the main study.

4.4.3.2. Improvements for the experimental design

Principally, the set-up of the first study is maintained for the main study to see if the unexpected results as in the contradicting measures present a systematic effect. Some of the changes concerned the experimental conditions as well as the materials, as a compromise concerning the required number of participants demanded a reduction of the experimental conditions. Because of the strong impact of the l-l condition for example shown in HR, this manipulation was excluded. It might interfere with the intention to provoke a state of monotony and thus affect the results in an unexpected manner. Also, to include this condition as control condition in the main experiment might have led to a significant main effect caused by this manipulation and thus reduced the impact of any additional effects imposed by the sequence of DD. The reason is that in the condition of low DD in two runs, physiological activation, effort and sleepiness was increasing, which might also indicate stress-related effects of counteracting underload. Even though this condition might have helped to better classify the results in relation to the theories explaining monotony, only two groups were maintained for the main study. Nevertheless it is recommended to investigate this effect in future studies.

The multiple measurement approach describing a specific temporal development of the indicators was maintained for the main study. Also, the duration of the scenarios was long enough to demonstrate the effects of repetitiveness. The non-linear development of some indicators needs to be considered when formulating the hypotheses, as interactions can be expected. As the impairments in the most important indicators were already reflected at the second measurement point, the scenario duration was reduced to 45 minutes. In subjective ratings, the 3 measurement points after level correction were maintained for the analysis to better compare the results of the small-scale and main simulation study.

To estimate the sample size needed for the main study, effect size (ES) was calculated for the main effects (Table A-19). These were calculated through consideration of all available measures for each indicator, independent on the repeated measurement variable. For the main effect repetitiveness, the effect size is very large (ES=2.08). Cohen's (1992) d' was used to calculate the effect size of the indicators for the repetitiveness and also for the sequence of DD. To use the between subjects value without correcting for the correlation between the two time periods is in line with arguments developed by Dunlop, Cortina, Vaslow, and Burke (1996) who stated that the correction would overstate the ES between two scores. Because of the decision to only include the sequence groups for the DD manipulation, also the effect size was only compared between the high and the low DD condition aggregated over the scenarios and resulted in a high effect of ES=.78. The determination of power was excluded in this small-scale experiment, since this was not designed to come to conclusions on the significance of the effects, but rather to test the set-up and gain an increased understanding of the effects on the main indicators.

4.4.3.3. The selection of materials

The variety of materials deployed in the small-scale experiment was used to assess contradictions and similarities on different levels and resulted in some unexpected results. For example, fatigue was reflected in individual ratings, but not in the fatigue subscale. This might be a consequence of the different measurement schedule, but also speak against the validity of the employed instruments, as the combination of the other physiological and subjective measures was supporting the assumption that fatigue actually occurred. These contradictions will need to be considered and request a careful interpretation of the results that will be obtained in the following studies.

The combination of reconstruction interviews as well as additional debriefings was taking too long. For this reason, in the main study, the final debriefing included the questions of the reconstruction interview. SASHA will be excluded as it did not contribute to enlighten the relationships concerning situation awareness and in its place, a one-item measurement of situation awareness is introduced, which also reduced the number of statistical tests.

4.5. SUMMARY

The goal of the experimental study was to test the hypothesis if effects of task characteristics, which were assumed to provoke monotony, were reflected in different psychophysiological indicators. In a small-scale laboratory set-up, eight former air traffic controllers executed two scenarios of 50 minutes each, which consisted of either repetitive or non repetitive traffic patterns. All combinations of high or low dynamic density (DD) were completely permutated during two runs. Physiological variables included heart rate (HR) and heart rate variability (HRV), skin conductance level (SCL) and blink rate. On a subjective level, cognitive, emotional and motivational aspects were assessed during the task and questionnaires for mood states and critical states were administered after each scenario. After-task performance was assessed with a test battery.

On a physiological level, especially the development of HR and SCL reflected decreasing physiological activation. The effect of dynamic density was reflected in HRV, but was mainly the consequence of the strong effect of DD in the low - low sequence in both scenarios. Subjectively, in the repetitive condition controllers felt sleepier and more fatigued as well as less attentive. Boredom and irritation were affected differently. A significant effect of repetitiveness was found in the feeling of monotony. Complex reaction patterns resulted in interactions with time-on-task factors. Unexpected effects were related to contradicting results in the subscales for mood and critical state assessment.

The findings indicate that the set-up was able to demonstrate the expected results and thus could be maintained for the main study. The results supported the multi-method approach as well as the selection of the strongest indicators for the state of monotony, which can be used in statistical analysis in the main study. Minor changes concerned the manipulated conditions for the main study, the length of the scenarios as well as the used questionnaires.

5. STUDY II: A SIMULATOR STUDY TO DETERMINE FACTORS EVOKING MONOTONY IN ATC

5.1. OBJECTIVES AND HYPOTHESES

The current study is a continuation of the small-scale study and addresses all presented objectives, which are the investigation of the effects of task characteristics (e.g., repetitiveness), additional contributing factors (e.g., individual factors) and countermeasures (e.g., variation in activity). The overall scope of the simulation approach is the description of a state of monotony under the controlled conditions of an experiment. Where the first experiment allowed defining the best indicators for a state of monotony and validating the set-up, the present study had the goal to determine relevant factors based on statistical decisions. On the basis of the previously obtained results, hypotheses about the effects of task characteristics on the state of monotony were tested in this study. In addition to the description of the psychophysiological developments, there was an interest in how repetitive traffic conditions affect the performance of controllers. Since a complex reaction pattern was expected on multiple levels, as shown in the first study, a further interest was the distinction of critical states indicated by various measures. According to the theory and previous results only certain indicators were assumed to reflect distinguished states, consequently very selective hypotheses were tested. Situational variables were held constant in the experiment, and individual variables included states as well as traits. Additional hypotheses for all other indicators were formed for a precise description of task effects in a complex reaction pattern.

The difference between descriptive and confirmatory study was already described in 2.3. As one study can have confirmative as well as descriptive hypotheses, it is important to indicate which type is dealt with. This was considered through indices. In the following, the confirmative and descriptive research hypotheses are summarized, which were tested in this study.

Main confirmative hypotheses:

Hypothesis II.1.1-2_C: There is a difference in the development of composed indicators for the state of monotony in air traffic controllers depending on the traffic characteristics repetitiveness and dynamic traffic density.

Hypothesis II.2.1-3_C: There is an influence of the individual characteristics boredom proneness and initial state of strain and recovery, on composed indicators for the state of monotony.

Hypothesis II.3_C: There is a difference in composed indicators for the state of monotony in a further run depending on prior break activity.

Additional descriptive hypotheses:

Hypothesis II.4D: There is a difference in the development of physiological, subjective and behavioral indicators depending on repetitiveness and sequence of DD.

Hypothesis II.5_D: There is a difference in the development of composed indicators for the critical states of monotony, satiation and fatigue.

5.2. METHOD

5.2.1. Experimental Design

The experimental design for the main study involved a 2 x 2 x 2 x (3 vs.15) mixed design (Table 24). Similar to the prestudy, it included the between-subject factor repetitiveness (non repetitive/repetitive conflict patterns in traffic), the partially permutated between-subject factor sequence of Dynamic Density (DD) (sequence from low to high (l-h) versus high to low(h-l)), the within-subjects factor run (first/second) and the within-subjects factor interval during run (i=3 vs.15). In addition to the first study, the between-subject factor activation in rest break was included. Again, participants were randomly assigned to the experimental conditions.

Table 24: Experimental design of study II

BREAK ACTIVITY	active				non active			
REPETITIVENESS	repetitive		non repetitive		repetitive		non repetitive	
	DD	DD	DD	DD	DD	DD	DD	DD
Run 1	l	h	l	h	l	h	l	h
Run 2	h	l	h	l	h	l	h	l
Break								
Run 3								
n	3	3	3	3	3	3	3	3
n total	24							

Note.(sequence of Dynamic density (DD): l=low, h=high)

5.2.1.1. Independent variables (IV)

Most of the factors included in the design have already been described in detail in chapter 4 and thus only differences are presented. To summarize, the following factors were included:

- Break activity: active vs. non active break.
- Repetitiveness: repetitive vs. non repetitive traffic.
- Sequence of Dynamic Density (DD): scenarios are changing from high– low vs. low – high DD in two runs (h-l: Run 1: high DD, Run 2: low DD; l-h: Run 1: low DD, Run 2: high DD).
- Run: first vs. second run.
- Interval: 3 (vs. 15) intervals during one scenario.

The additional factor break activity was realized through the introduction of short rest breaks of ten minutes after finishing the second simulation run. In the non active condition controllers were instructed to spend the break in a relaxed position on a chair and close their eyes. For the active break they were instructed to execute physical exercises. The physical exercises consisted in seven selected isometric and stretching exercises for the neck, shoulders and activating the cardiovascular system. They were each based on the principle of alternating tension and relaxation which lasted for around 15 seconds each time. It was possible to execute all of them whilst seated. Instructions were based on ones that have been used in an earlier study (Straussberger & Kallus, 2003) on rest break design. A deviation occurred in the factor of the sequence of Dynamic Density (DD), which was only manipulated in two conditions. It is emphasized that the sequence of DD was chosen to determine the effect of high and low DD through significant interaction effects in trend components. In addition, due to the reduced duration of the scenarios the intervals during scenarios contained only 15 periods.

5.2.1.2. Dependent variables (DV)

Dependent variables were assessed considering the multi-level approach including physiological, subjective, and behavioral measures. Its advantage is to develop a complex profile of effects.

The composition of an indicator for the state of monotony

In the prestudy it was observed that a state of monotony caused by task repetitiveness is reflected in the subjective feeling of monotony, sleepiness and heart rate. This reflects the results of Bartenwerfer (1957) and Thackray et al. (1975) and thus corresponds with the theoretical assumptions. For the purpose of testing the main hypothesis, these indicators were submitted to the MANOVA to explain the overall contribution. Thus it deviates from the recommended procedure by Miyake (2001), who created a single workload indicator with combined subjective and physiological measures, but was seen as necessary because of the weak internal consistency. To achieve a consistent pattern of activation, the polarity of heart rate was changed when submitted to MANOVA. Finally, all variables that were included in this indicator were standardized into a z-score to provide comparable units of measurement in different variables. Even though in original theoretical descriptions of monotony performance impairments were a major descriptor, this aspect was not included in this indicator. As known, performance can be maintained on a high level, although an individual already feels changes on a subjective level. It could be assumed that performance is only affected in a very late stage of monotony. For this reason performance was not chosen as a main criterion for the rise of a state of monotony.

Physiological measures

Additional physiological indicators from the cardiovascular and peripheral system were collected. In addition to the measures of the prestudy, brain activity was analyzed. Table 25 shows a summary of physiological measures.

Table 25: Summary of physiological variables (Study II)

INDICATOR	DEPENDENT VARIABLE
ECG	Average heart rate (corrected for baseline) in 3-minute-intervals during run
	Average heart rate (corrected for baseline) in performance tests
	Average heart rate variability in 3-minute-intervals during run
	Average heart rate variability in performance tests
EEG	Average power in beta (corrected for baseline) in CZ in 3-minute-intervals during run Average power in theta (corrected for baseline) in CZ in 3-minute-intervals during run Average power in alpha1 (corrected for baseline) in CZ in 3-minute-intervals during run Average power in alpha2 (corrected for baseline) in CZ in 3-minute-intervals during run
EDA	Average skin conductance level (corrected) in 3-minute-intervals during run
	Average skin conductance level (corrected) in performance tests
EOG	Average number of blinks in 3-minute-intervals during run

Subjective measures

The collection of subjective indicators took place before, during and after each run. Table 26 shows a summary of the subjective indicators and their operationalization.

Table 26: Summary of psychological variables (Study II)

INDICATOR	DEPENDENT VARIABLE
Cognitive, emotional and motivational indicators (TSI)	Average scores of *attentiveness* (level-corrected) at 3 points of measurement during run
	Average scores of *fatigue* (level-corrected) at 3 measurement points during run
	Average scores of *boredom* (level-corrected) at 3 measurement points during run
	Average scores of *irritation* (level-corrected) at 3 measurement points during run
	Average scores of *strain* (level-corrected) at 3 measurement points during run
	Average scores of *concentration* (level-corrected) at 3 measurement points during run
	Average scores of *motivation* (level-corrected) at 3 measurement points during run
	Average scores of *sleepiness* (level-corrected) at 3 measurement points during run
Mood (UWIST)	Average scores in *hedonic tone* after each run
	Average scores in *tense arousal* after each run
	Average scores in *energetic arousal* after each run
Critical States (SOF)	Average scores in *stress* after each run
	Average scores in *fatigue* after each run
	Average scores in *monotony* after each run
	Average scores in *satiation* after each run
Workload (NASA-TLX)	Average scores in *mental demand* after each run
	Average scores in *physical demand* after each run
	Average scores in *temporal demand* after each run
	Average scores in *frustration* after each run
	Average scores in *effort* after each run
	Average scores in *performance* after each run
	Average scores in *overall workload* after each run
Additional Indicators	Average scores in *feeling of monotony* after each run
	Average scores in *situation awareness* after each run

They indicated the controller's interpretation of the situation, and the psychological state concerning his or her mood and well-being.

Behavioral and performance measures

In the main study a more precise evaluation of performance indicators connected with a state of monotony was undertaken (Table 27).

For performance measures it was considered that the primary task of controllers is to provide sufficient separation between aircraft. As such STCA alerts are the clearest indicator of reduced performance. But this is an event that usually does not occur that often. Therefore, it was reasonable to assume that effects on performance would not necessarily be reflected in commonly used performance measures in ATC. Nevertheless to have an indication for this situation it was interesting how much in advance a potential conflict was noticed. This indicator was thus similar to the one used by Athenes, Colles, and Dittmar (2002) to describe mental workload. The authors defined Maturing Time (MT) as the interval of time between potential conflict diagnosis and the moment when a conflict can be regarded to no longer exist. The idea of introducing an unexpected scenario was also due to the result of Dougherty, Grondlund, and Durso (1999) who found in 12 ATCOs that they rather generate a single satisfactory plan instead of different plans and then make that plan work. This might be a further factor contributing to the manifestation of monotony and provide evidence for the set-effects of Luchins (1942), as it is not efficient for an ATCO to look for additional solutions, once a solution was generated.

In consequence, different behaviors might be executed to counteract a suboptimal state and thus needed to be asked for. The use of tools might be more frequent in underloaded monotonous conditions to keep oneself occupied. But it contains the risk of gathering inappropriate information with the potential to miss important information. Again, the use of additional performance tasks after the run of the experiment was introduced. Other reasons concerning the announcement of suboptimal states or to increase activation have been discussed by Rogé, Pebayle, and Muzet (2001) to explain behavioral activities. It was expected that performance impairments were not visible at first sight, as they might be announced through small mistakes such as late acceptance or hand-over, giving wrong inputs, especially with increasing time on task. Information on these aspects were collected in the debriefing.

Table 27: Summary of performance variables (Study II)

INDICATOR	DEPENDENT VARIABLE
Primary task performance	No. of STCA alerts in extraordinary potential conflict situation
	Relative mean time for conflict solution in extraordinary potential conflict situation
After Task performance	Vienna Reaction Test: Average scores in motor time
	Vienna Reaction Test: Average scores in reaction time
	Cognitrone: Average scores in reaction time to correct answers
	Time-Movement-Anticipation Test: Median of total deviation time in sec
	Time-Movement-Anticipation Test: Median of total direction deviation in pixels

5.2.1.3. Moderator and control variables

Potential moderator variables have already been included for control purposes in the first study and are summarized in Table 28. Some of these indicators were expected to influence the dependent variables and thus were considered in blocking designs or analysis of covariance. A number of control variables were collected to ease the interpretation of significant differences between groups: initial state of the subjects (ISQ), body movements, age, expertise, gender, handedness, body weight, body height, smoking, time, and room temperature. Due to organizational reasons two experimental sessions took place at one day, starting at 8 AM and 14 PM respectively. The starting time was counterbalanced between the conditions.

Table 28: Summary of potential moderator variables and applied scales (Study II)

INDICATOR	VARIABLE
Initial states of recovery and strain (RESTQ)	Average scores in aggregated *recovery* subscales
	Average scores in aggregated *stress* subscales
	Average scores in subscales (General Stress, Emotional Stress, Social Strain, Conflicts, Overfatigue, Lack of Energy, Somatic Complaints, Success, Social Recovery, Somatic Recovery, General Recovery, Recovery Sleep)
Action Control Style (ACS)	Average scores in *Decision-related Action Orientation* (AOD)
	Average scores in *Action Orientation after Failure* (AOF)
	Average scores in Action Orientation during Successful Performance (AOP)
Morningness-eveningness-preference (MEQ)	Average scores in morningness-eveningness preference
Boredom proneness (BPS)	Average score in boredom proneness
Personality domains (IPIP)	Average scores in *extraversion*
	Average scores in *agreeableness*
	Average scores in *conscientiousness*
	Average scores in *emotional stability*
	Average scores in *intellect*

5.2.2. Procedure

The experiment was conducted between the 7th and 30th June 2004 at the premises of the Maastricht Upper Area Control Centre (MUAC). With the support of the rostering service an invitation leaf-let was sent out in the ops room environment to recruit controllers for the study. The volunteers were released from work to participate in the experiment and received a controller handbook before they were scheduled for participation. Minor changes in the handbook compared with the pre-study concerned the mentioned changes in the materials and procedure. The study consent form was completed before the experimental session took place and participants were asked not to drink caffeine drinks or smoke during the session. Because of the long duration of the experimental session, local requirements such as a free break of 20 minutes needed to be considered in the set-up, which was implemented after the familiarization phase.

Table 29: Experimental procedure (Study II)

STEP	TIME IN MIN	TIME CUM
PREPARATION		
Welcome and Briefing	10	10
Preparation of physiological measures + Questionnaires (ISQ, RESTQ)	40	50
Training eDEP	30	80
Training VTS (Vienna Test System)	10	90
Free Break	20	110
MEASURED RUNS SCENARIO 1		
Baseline + TSI	4	114
Scenario 1 (incl. TSI every 15 min)	45	159
Baseline	3	162
Scales (UWIST, SOF, NASA)	7	169
MEASURED RUNS SCENARIO 2		
Baseline + TSI	4	173
Scenario 2 (incl. TSI every 15 min)	45	218
Baseline	3	221
Scales (UWIST, SOF, NASA)	7	228
MEASURED RUNS SCENARIO 3 WITH BREAK INTERVENTION		
Short Break with Activity/No Activity	10	238
Baseline + TSI	3	241
Scenario 3	15	256
Baseline	3	259
Scales (TSI, UWIST, NASA)	4	263
Test (RT, COG, ZBA)	15	278
AFTER RUNS		
Remove electrodes	5	283
Debriefing and Discussion	20	303
TOTAL in min	303	

The simulation environment was set up in a room provided by the control center, with an average temperature of 26 C (+/- 1 C). This temperature was rather high because of local restrictions and extreme weather conditions in this period. Table 29 provides an overview about the experimental procedure, which was conducted similar to the preliminary study. One exception is the length of the first two scenarios reduced to 45 minutes. In addition, a third scenario of 15 minutes was introduced after the short break to evaluate break effects and questionnaires were again administered before and after the scenario. On the average one experimental session lasted for 312 minutes (SD=31).

5.2.3. Participants

In this study, the parameters to assess the required sample size for sufficient power of 0.80 were set to alpha=0.05 and beta=0.2. For repeated measures, correlations with an average of 0.5 were assumed. Power estimations were only conducted for the main factors of interest that is repetitiveness, break activity and the sequence of dynamic density. The power for detecting an effect of DD was determined for the interaction between sequence of DD and run, which was used to determine differences in the impact of high or low DD. According to the effect-size calculations in the small-scale study, a strong effect could be expected for group differences. For a statistical power of 0.80, a minimum sample size of 22 was required to detect a truly significant effect in the main factors. This was rounded up to 24 to provide equal cell sizes in the experimental conditions. The detailed approach is described in Appendix B.1.1 and is based on Cohen (1992). Calculations were undertaken with *G-Power* (Erdfelder, Faul, & Buchner, 1996; see Appendix B.1.1).

The participants were recruited through the help of the local rostering office, which was asked to choose fully qualified volunteers, with an age preferably between 20 and 40 years. Also, the common ratio of male to female in the center of 4:1 should to be maintained.

Twenty-four volunteering ATCOs from ten different European nations participated in this study (Table B-1). One additional ATCO contributed in an initial trial session, which was introduced to test the set-up and procedure. 21 ATCOs were licensed for enroute, 3 had additional ratings for tower and/or approach control. Fifteen of the controllers were also instructors. One controller was more than 40 years old. Further descriptive statistics and frequencies are presented in Table 30.

Table 30: Sample statistics and frequencies (Study II)

	MIN	MAX	M	SD
Age in years	21	47	29.5	6.1
License in years	0.2	21.0	6.1	5.6
Height in cm[a]	164	202	179.7	8.2
Weight in kg [a]	58	102	77.9	12.6
	FREQUENCY	n		
Gender	male	18		
	female	6		
Handedness	left	3		
	right	20		
	both	1		
Vision [a]	normal	18		
	glasses	5		

Note: N=24; [a]missing: n=1

There are no significant group differences in the initial state of strain or recovery, the experience or the age (Table B-2).

5.2.4. Materials and Apparatus

Table 31 gives an overview of the used questionnaires and scales. They have already been described in detail in chapter 4.2. Additional materials and deviations from the previously administered ones are reported in the following.

Table 31: Summary of materials and apparatus

ADMINISTRATION	NAME (ABBR.)	AUTHOR
Before session	Biographic questionnaire (BIO)	
	Personality Inventory (IPIP)	Goldberg 1999a
	Morningness Eveningness Questionnaire (MEQ)	Horne and Ostberg, 1976
	Action Control Style (ACS 90)	Kuhl, 1994b
	Boredom Proneness Scale (BPS)	Farmer and Sundberg, 1986
Beginning of session	Initial State Questionnaire (ISQ)	Translation based on Janke, 1976
	Questionnaire for Strain and Recovery State (RESTQ)	Kallus, 1995
During session	Workload (NASA TLX)	Hart & Staveland, 1987
	Thackray Scale Inventory (TSI)	Based on Thackray, 1975; Barmack, 1939
	Scale of feelings SOF-II	Translated version of BMS (Richter & Plath, 1984) by Rockstuhl, 2002;
	Mood Assessment UWIST	Matthews, Jones & Chamberlain, 1990
End of session	Debriefing guide	

The questionnaire addressing the initial state (ISQ) was completed with an item asking for the work schedule of the last three days. The questionnaire to assess situation awareness (SASHA) was removed. Instead, one item to rate situation awareness was included in the NASA-TLX. The interviews already described in the preliminary study were only conducted at the end of the session and thus the debriefing questions were adapted.

Personality Inventory (IPIP)

This freely available personality inventory is based on Goldbergs (1999a) International Personality Item Pool and was used to assess the influence of personality domains. It is based on the currently dominant model that personality can be described in terms of a hierarchical model with five areas (Goldberg, 1990; Costa & McCrae, 1992; Digman,1990), often referred to as Big Five Factors. The instrument was chosen because it addresses the same constructs as the NEO-FFI, (the questionnaire commercially published but for the study not available), is short, and Cronbach's Alpha has been reported sufficiently high (Table 32). Each subscale consists of 10 summed up items and response options range from 1 (very inaccurate) to 5 (very accurate).

Table 32: Overview and description of personality inventory (IPIP) subscale

SUBSCALE	ABBREVIATION	EXAMPLE ITEM	DESCRIPTION	α^a
Extraversion	extra	I start conversations.	Preference for and behavior in social situations	.87
Agreeableness	agree	I am interested in people.	Interaction with others in terms of trusting, friendly and cooperative opposed to aggressive	.82
Conscientiousness	cons	I am always prepared.	How organized and persistent one pursues his or her goals	.79
Emotional stability	emot	I get upset easily.	Tendency to experience negative thoughts and feelings	.86
Intellect	intell	I have excellent ideas.	Open-mindedness and interest in culture	.84

Note. [a] Cronbach's Alpha

5.2.5. Data Processing

The data processing was already described in 4.2.5. This chapter only contains additional procedures or deviations. The data correction method used was the same as in the prestudy: HR was corrected for the baseline-values (4 baselines collected) and uncorrected HRV was used, which was supported by the fact that in the first three measurement periods no difference between the two task conditions was found.

Additionally, in the main study EEG was analyzed using *Vision Analyzer* 5.1 (Brain Products GmbH, Germany). The EEG recordings were examined for artifacts, poor recordings or blinks. The following procedure was applied as recommended by the provider. Blinks were corrected using the regression approach proposed by Gratton and Coles. Raw recordings were visually checked for inactive periods. The segmenting of the EEG was done with respect to the onset of the traffic scenarios in 3-minute-intervals. Segments with large EEG changes were detected (if changes within one segment > 100 mV; adapted to individual differences if necessary) and eliminated by the program. Finally, absolute power was computed for each frequency band which were divided into theta (5-7 Hz), alpha 1 (8-10 Hz), alpha 2 (11-13 Hz), and beta 1 (13-20 Hz). Baseline-corrected values were submitted to statistical analysis. Concerning the input of the performance data as well as the subjective data the same procedure was applied as in the small-scale-study. All data was imported in SPSS 14, were statistical procedures were executed.

5.2.6. Statistical Analysis and Hypothesis Testing

The initially formulated hypotheses addressed the effect of task characteristics on monotony and other psychophysiological indicators. These research hypotheses were translated into a set of statistical hypotheses and submitted to statistical test procedures. In general, the null hypothesis for both the main and interaction effects assumes that the variance of the effect and the variance of the residuals do not differ.

Each confirmative hypothesis is submitted to the appropriate statistical analysis with repetitiveness and sequence of DD as between-subject factors and run as within-subject factor in Hypothesis II.1. The confirmative hypotheses were tested with a multivariate analysis of variance (MANOVA) including the three indicators of monotony as dependent variables. The advantage of this procedure was that it allowed not only the determination of the overall effect, but also the interpretation of individual indicators. A second confirmative hypothesis addressed the influence of the moderator variables initial states of strain and recovery and boredom proneness. Depending on the preconditions, the analysis of covariance or a blocking design was used to test the effects. The final confirmative hypothesis tested the effect of break activity and repetitiveness on monotony. For confirmative hypotheses, an overall alpha level of .05 was used for statistical tests. Since the variables were not independent, the alpha-level was adjusted according to the sequential Bonferroni-Holm procedure. The p-values were put in increasing order and the smallest p-value had to remain under alpha/*k*, where *k* was the total number of hypotheses tested. The second smallest p-value was lower than alpha/*k*-1 to reject a null hypothesis. The alpha-level of .05 was maintained as the results were the basis for recommendations for future concept developments. Bonferroni-Holm was the method of choice, as it avoided alpha inflation without being too conservative. It was also superior to procedures like the creation of hypotheses families, as it did not require a decision on the indicators to be rejected by the null-hypotheses, but split up the level of alpha on all available tests. Finally, a method like the evaluation of an overall F-value was not appropriate, as because of the different temporal development of the indicators a different amount of the error variances might have been attributed to interaction terms in the combined indicators.

The analysis of covariance was undertaken to investigate the moderator variables. This requires however that there is a significant correlation between the dependent variable and the moderator variable. If this condition was not satisfied, the moderator variable was blocked and introduced in the design as a third independent variable.

Differences in the development of effects of DD were determined from trend analysis. It is noted that any significant interactions between sequence of DD and run expressed the effect of counterbalancing. Therefore, trend analysis for description of the dynamic density was used to determine its effect and the trends are also described for the course of the indicators over time. As such, a significant linear trend in the interaction between DD and Run demonstrates an effect of the DD and can be further described through the comparison of descriptive statistics.

The interpretations of main effects of repetitiveness were based on the assumption that there were no significant interactions with sequence of DD. In case where not stated differently, insignificant interactions between variables were seen as an additional precondition to accept H1. However, the type of the interaction needs to be considered for their evaluation. Any interactions with run and intervals did not affect the interpretation of the main effects, as long as the main effect was reflected throughout the time. A different development during each run was expected because of the manipulation of the independent variable and rather the direction of the size was determined.

For confirmative hypotheses the assumptions of normally distributed variables for the subgroups of each factor condition were proven with Kolmogorov-Smirnov-Test and equality of variances with Levene's Test. In case the assumption of sphericity was not met as revealed through a significant result in Mauchly's test, degrees of freedom (*df*) were corrected according to the method of Greenhouse-Geisser.

Additionally, descriptive hypotheses addressing the effects of the task characteristics on additional indicators and statistical analysis employed a full model repeated measure ANOVA applying the general linear model with repetitiveness and sequence of DD as between factors, and run and intervals within run as within factors. According to the manipulation described in 5.2, the dependent variables were included for different levels. The procedure was identical to the one in the first study. For nominal variables, an exact Fisher Test was used to determine differences between experimental conditions. A note concerns the interpretation of interactions, where several possibilities were discussed by Petty, Fabrigar, Wegener, and Priester (1996). The interpretation of trend effects was preferred concerning the time course of the indicators and simple effects were calculated if necessary. For the purpose of convenience, the descriptive statistics and results of the statistical analysis are reported in Appendix B if not stated otherwise.

5.3. RESULTS

5.3.1. Confirmative Hypotheses 1: Effects of Task Factors

The major hypothesis in this experiment was that the task factors would have an effect on a state of monotony. For monotony, a combination of indicators was chosen. To adapt heart rate (HR) to the other indicators, it needed to be inverted[9], which needs to be considered in the tables and graphs, where an increase in the value means lower HR. Table 33 shows the mean values and SD for each group of indicators

[9] If inverted HR is presented, it is clearly referred to in figures and tables.

Table 33: Average values (and SD) for the indicators of monotony for each run depending on repetitiveness and sequence of DD (Study II)

		Repetitive		Non repetitive	
	DD	**l-h**	**h-l**	**l-h**	**h-l**
HR (inv)	Run 1	-3.87 (1.16)	-4.59 (3.85)	-6.41 (3.42)	-4.27 (1.25)
	Run 2	-1.91 (2.35)	-1.54 (2.63)	-5.72 (1.81)	-2.84 (1.84)
Sleepiness	Run 1	.10 (.77)	-.56 (.41)	-.28 (.38)	-.63 (.66)
	Run 2	.32 (.84)	1.17 (.57)	-.11 (.71)	.59 (.32)
Feeling of monotony	Run 1	83.83 (12.72)	53.17 (13.03)	37.67 (26.47)	44.33 (19.62)
	Run 2	46.50 (21.95)	68.83 (20.14)	59.50 (30.63)	57.33 (21.83)

Note. N=24. Sequence of dynamic density (DD): l-h: low DD in Run 1, high DD in Run 2; h-l: high DD in Run 1, low DD in Run 2.

The indicators were submitted to a 2 x 2 x 2 MANOVA with repeated measures on the last factor. All included indicators were normally distributed (all p>.598).

The multivariate test revealed a significant main effect of repetitiveness (Wilks' Lambda=0.48, $F_{3,18}$=6.451, p=.004; η_p^2=.52) and a significant main effect of run (Wilks' Lambda=0.407, $F_{3,18}$=8.754, p=.001; η_p^2=.59). No significant effect of sequence of DD was found (Wilks' Lambda=.82, $F_{3,18}$=1.295, p=.307; η_p^2=.18). In addition, significant two-way-interactions between run and repetitiveness (Wilks' Lambda=0.64, $F_{3,18}$=3.379, p=.041; η_p^2=.36) and run and sequence of dynamic density (Wilks' Lambda=0.60, $F_{3,18}$=4.066, p=.023; η_p^2=.40) were found, no significant interaction between repetitiveness and sequence of DD resulted (Wilks' Lambda=.88, $F_{3,18}$=.850, p=.485; η_p^2=.12). Finally, also the 3-way-interaction between run, repetitiveness and sequence of DD was significant (Wilks' Lambda=0.65, $F_{3,18}$=3.282, p=.045 η_p^2=.35).

To determine the magnitude of the linear trend component, also the composed standardized indicator for the state of monotony was submitted to a statistical test and revealed a significant linear trend in the sequence of DD x Run interaction ($F_{1,20}$=12.93, p=.002).

As can be seen in Table 34, the null-hypotheses for the effects of repetitiveness and dynamic density can be rejected and the alternative hypotheses assumed.

To further describe the results, the univariate F-Tests were examined for each variable to identify which specific dependent variable contributed to the significant overall effect. Mean and SD are presented in Table 33, the results of the statistical analysis can be found in Table 34. Inverted HR in run 1 did not reveal equally spread variances ($F_{3, 20}$=4.41; p=.015), but the sphericity assumption was met for all indicators. A significant effect of repetitiveness was found in sleepiness, but the effect was only tendencially significant in heart rate (inv.) and the subjective feeling of monotony. The significant time effect indicated that heart rate was lower and sleepiness greater in the second run compared to the first run.

Table 34: Results of univariate Analysis of Variance for indicators of monotony (Study II)

Source of Variance	Results ($F_{df\ hypothesis,\ df\ error}$, p-value)		
	HR inv (corr. baseline)	Sleepiness	Feeling of Monotony
REP	***$F_{1,20}$=4.01, p=.059(*)*** [a]	**$F_{1,20}$=10.01, p=.005*** [a]	***$F_{1,20}$=3.26, p=.086(*)***
DD	$F_{1,20}$=1.63, p=.216	$F_{1,20}$=1.32, p=.263	$F_{1,20}$=0.01, p=.898
RUN	**$F_{1,20}$=18.67, p=.000***	**$F_{1,20}$=14.35, p=.001***	$F_{1,20}$=0.48, p=.497
Rep x DD	$F_{1,20}$=2.15, p=.158	$F_{1,20}$=.11, p=.741	$F_{1,20}$=.19, p=.670
Run x Rep	***$F_{1,20}$=3.05, p=.096(*)***	$F_{1,20}$=0.40, p=.535	**$F_{1,20}$=8.83, p=.008*** [a]
Run x DD	$F_{1,20}$=1.22, p=.282	**$F_{1,20}$=8.43, p=.009*** [a]	**$F_{1,20}$=5.39, p=.031*** [a]
Run x Rep X DD	$F_{1,20}$=.05, p=.833	$F_{1,20}$=.26, p=.619	**$F_{1,20}$=10.57, p=.004*** [a]

Note. N=24. Rep=Repetitiveness; DD=Sequence of DD. Trend effects: [a]linear. ***p<.001. **p<.01. *p<.05. (*)p<0.1.

Figure 23 depicts the development of sleepiness. As the inverted HR (inv.) showed a similar course, it was not depicted in a graph. On the other hand, a different development is reflected in the feeling of monotony.

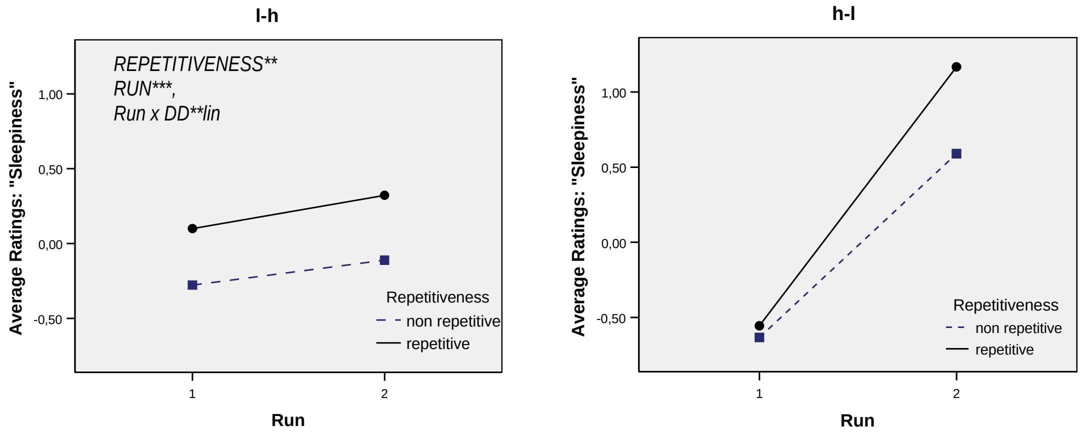


Figure 23: Average ratings of sleepiness for each run as a function of repetitiveness and sequence of dynamic density (DD) (Legend: l-h: low DD in Run 1, high DD in Run 2; h-l=high DD in Run 1, low DD in Run 2 (Legend: ● black continuous line: repetitive; ■blue dotted line: non repetitive).

The significant interaction between run and sequence of DD reflects the effects of the treatment and was significant in the linear trend component in sleepiness and feeling of monotony. The higher order interaction in the subjective feeling of monotony is displayed in Figure 24. Further comparisons revealed that the subjective feeling of monotony develops differently depending on the sequence of DD. Feeling of monotony was rated higher in the first run of the repetitive traffic under low DD ($F_{1,20}$=18.05, p=.000; η_p^2=.31). HR was more deactivated under the repetitive compared to the non repetitive traffic condition of the second run under high DD ($F_{1,20}$=9.14, p=.007; η_p^2=.31).

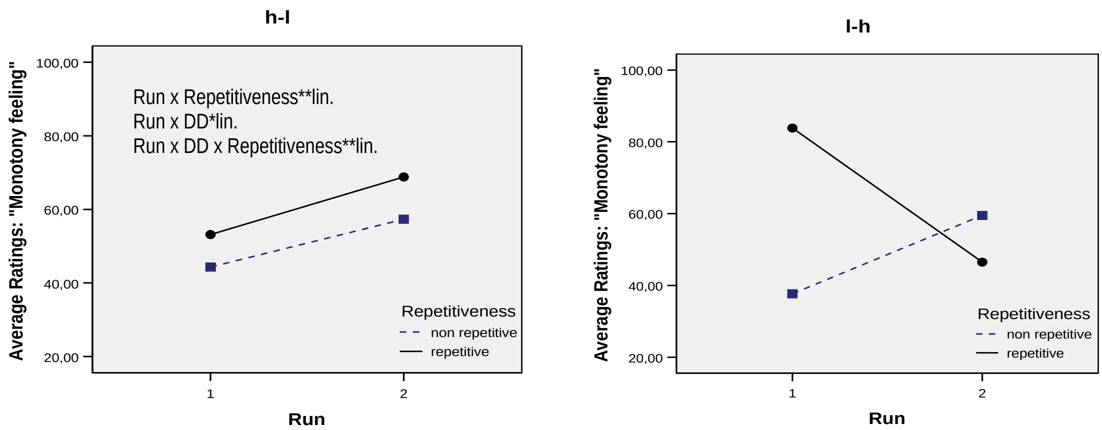


Figure 24: The average ratings of the subjective feeling of monotony for each run as a function or repetitiveness and sequence of dynamic density (DD) (Legend: l-h: low DD in Run 1, high DD in Run 2; h-l=high DD in Run 1, low DD in Run 2 (Legend: ● black continuous line: repetitive; ■blue dotted line: non repetitive).

5.3.2. Confirmative Hypotheses 2: Effects of Moderator Variables

To investigate the influence of initial state of strain, recovery and boredom proneness, MANOVAs were performed with median-split variables as additional factor. Descriptive statistics and detailed results are depicted in Table B-3 and Table B-4.

There are no significant main effects of boredom proneness, initial state of recovery and strain on the indicator for the state of monotony (main effects: all p >.447; Table B-4). Also, no significant interactions between these and traffic-related factors were found. As can be seen in Table 35, the null-hypotheses that there is no effect of the mentioned moderator variables are maintained.

5.3.3. Confirmative Hypotheses 3: Effects of Countermeasures

The current study considered activity in rest breaks as one potential countermeasure to reduce the unwanted side-effects related to monotony. To investigate effects of break activity on the indicators of the state of monotony during the execution of a third scenario, MANOVA was conducted with the manipulation of break activity, repetitiveness and sequence of DD during the two scenarios as between-subjects variables. The submitted indicators were normally distributed (all p>.687) and variances were homogeneous (p>.148). It was hypothesized that no impact of break activity on monotony was observed depending on prior conditions. The results revealed a significant main effect of break activity (Wilks' Lambda=0.449, $F_{3,14}$=5.72, p=.009, η_p^2=.551) and a significant main effect of repetitiveness (Wilks' Lambda=0.347, $F_{3,14}$=8.79, p=.002, η_p^2=.653). No other source of variance was significant (all p>.223; descriptive statistics and further results in Table B-5 and B-6). Thus, again as presented in Table 35, the null-hypothesis for the effect of break-activity can be rejected and the alternative hypotheses assumed.

To further describe the results, the univariate F-Tests were examined for each variable to identify which specific dependent variable contributed to the significant overall effect. A significant effect of activity and repetitiveness was found in sleepiness (both effects: $F_{1,16}$=18.08, p=.001, η_p^2=.531), indicating lower values in the active break condition as well as in the repetitive condition. Heart rate (inv.) was also reduced in the repetitive condition (Repetitiveness: $F_{1,16}$=8.75 p=.009, η_p^2=.354), reflecting increased activation. The remaining effects on heart rate (inv.) and the subjective feeling of monotony were not significant (all p>.185).

5.3.4. Testing of Confirmative Hypotheses

For final decisions concerning the maintenance or rejection of the null hypothesis, the p-values were corrected according to the Bonferroni-Holm procedure. In a first step, all p-values were ranked starting from the smallest one. Table 35 enumerates the results of the statistical tests for the confirmative hypotheses and the decisions after alpha correction.

Table 35: Correction of alpha level for the confirmative hypotheses and related decisions (Study II)

Confirmative Hypothesis	Description	p-value	Rank	Adjusted alpha level	Decision for H0	Decision for H1
H1.1	Main effect of repetitiveness on monotony	.004	2	0.033:2=0.017	rejected	assumed
H1.2	Linear Trend in run x DD for monotony	.002	1	0.05:3=0.017	rejected	assumed
H2.1	Influence of boredom proneness	>.050	-	-	retained	rejected
H2.2	Influence of state of strain	>.050	-	-	retained	rejected
H2.3	Influence of state of recovery	>.050	-	-	retained	rejected
H3.1	Main effect of break activity on monotony	.009	3	0.026	rejected	assumed

After alpha correction, the null hypotheses addressing the task effects and the break effects were rejected; the null hypothesis addressing the influence of the individual factors were maintained. To further evaluate if the alternative hypothesis can be assumed, the significant interaction effects were assessed.

Concerning the effects of repetitiveness the alternative hypothesis can be assumed. The significant 3-ways interaction is mainly due to an expected impact of the run x sequence of DD interaction as a consequence of the DD manipulation. Also, the effect of DD could be confirmed. There was a different development of monotony depending on DD.

The results also supported the alternative hypotheses concerning the effect of break activity assumed, where no interaction effects needed to be considered in the evaluation of the hypotheses.

The null hypotheses for the effect of initial states of strain and recovery and boredom proneness were maintained.

5.3.5. Description of Additional Results

5.3.5.1. Effects of task characteristics on the physiological level

Physiological indicators were submitted to repeated measurement ANOVAs based on the general linear model. Because sphericity was not given for all repeated measures, Greenhouse-Geisser's corrected degrees of freedom were used. Levene's Test revealed heterogeneous variances in 10 percent of the HR variables, 25 percent of the HRV variables, and 10 percent of the SCL intervals. The indicators of two intervals during the second run were not normally distributed in HRV. The results of statistical analysis are presented in Table 36, descriptive statistics are presented in Appendix B (Table B-7 – Table B-10).

On the physiological level a significant main effect of repetitiveness on heart rate and heart rate variability was found. Heart rate was lower and heart rate variability higher in the condition of repetitive traffic, HR decreased and HRV increased from the first to the second run. No main effects were found for DD. This confirmed the aspect of physiological deactivation.

Figure 25 shows the development of heart rate during the first and second scenario, and Figure 26 represents the course of heart rate variability.

Table 36: ***Results of univariate Analysis of Variance for physiological indicators (Study II)***

Source of Variance	Results ($F_{df\ hypothesis,\ df\ error}$, p-value)			
	HR (corr. baseline)	HRV	No of. Blinks	SCL (corr.)
REP	$F_{1,19}$=4.38, p=.050*[a]	$F_{1,18}$=7.52, p=.013*	$F_{1,17}$=1.42, p=.250	$F_{1,11}$=.23, p=.644
DD	$F_{1,19}$=1.14, p=.298	$F_{1,18}$=.36, p=.558	$F_{1,17}$=2.37, p=.142	$F_{1,11}$=1.89, p=.197
RUN	$F_{1,19}$=20.38, p=.000***	$F_{1,18}$=24.98, p=.000***[a]	$F_{1,17}$=11.85, p=.003** [a]	$F_{1,11}$=1.27, p=.283
INT	$F_{5.6,105}$=1.03, p=.406 [g]	*$F_{6.4,117}$=2.12, p=.064* [g]	$F_{5.8,99}$=3.55, p=.003**[a]	$F_{1.9,21}$=2.13, p=.145[d]
Rep x DD	$F_{1,19}$=2.49, p=.131	$F_{1,18}$=1.91, p=.184	$F_{1,17}$=5.60, p=.030*	$F_{1,11}$=.17, p=.688
Run x Rep	*$F_{1,19}$=3.91, p=.063(*)*	$F_{1,18}$=2.51, p=.130	$F_{1,17}$=.00, p=.949	$F_{1,11}$=.03, p=.864
Run x DD	$F_{1,19}$=.69, p=.415	$F_{1,18}$=.00, p=.958	$F_{1,17}$=1.85, p=.192	$F_{1,11}$=.50, p=.496
Inter x Rep	$F_{5.6,105}$=1.19, p=.320	*$F_{6.4,117}$=2.05, p=.059(*)* [f]	$F_{5.8,99}$=1.82, p=.105[a]	$F_{1.9,21}$=.29, p=.741
Inter x DD	$F_{5.6,105}$=.65, p=.676 [e]	$F_{6.4,117}$=.87, p=.595 [g]	$F_{5.8,99}$=.84, p=.538	$F_{1.9,21}$=1.31, p=.290 [h]
Run x Int	$F_{4.8,91}$=.65, p=.657[h]	*$F_{5.2,93}$=2.11, p=.069(*)* [a]	$F_{6.2,106}$=1.97, p=.073[g]	$F_{3.3,37}$=3.52, p=.021* [b]
Run x Int x Rep	$F_{4.8,91}$=.45, p=.804	$F_{5.2,93}$=1.49, p=.199 [b]	$F_{6.2,106}$=.82, p=.564[c]	$F_{3.3,37}$=.28, p=.855
Run x Int x DD	$F_{4.8,91}$=.91, p=.476[h]	$F_{5.2,93}$=1.78, p=.122 [h]	$F_{6.2,106}$=.78, p=.595	$F_{3.3,37}$=.37, p=.796
Run x Rep x DD	$F_{1,19}$=.00, p=.963	$F_{1,18}$=.09, p=.772 [b]	$F_{1,17}$=8.48, p=.010*[a]	$F_{1,11}$=.72, p=.415
Int x Rep x DD	$F_{5.6,105}$=98, p=.436	$F_{6.4,117}$=1.23, p=.297	$F_{5.8,99}$=1.22, p=.302	$F_{1.9,21}$=1.1, p=.350
Run x Int x Rep x DD	$F_{4.8,91}$=1.51, p=.196	$F_{5.2,93}$=1.76, p=.126 [h]	$F_{6.2,106}$=.56, p=.765	$F_{3.3,37}$=1.19, p=.328

Note. N=24 Rep=Repetitiveness; DD=Sequence of DD; Int=Interval. Trend effects: [a]linear; [b]quartic; [c]cubic; [d]quartic; [e]order 5; [f]order 7; [g]order 8; [h]order 10; [i]order 14. ***p<.001. **p<.01. *p<.05. (*)p<0.1.

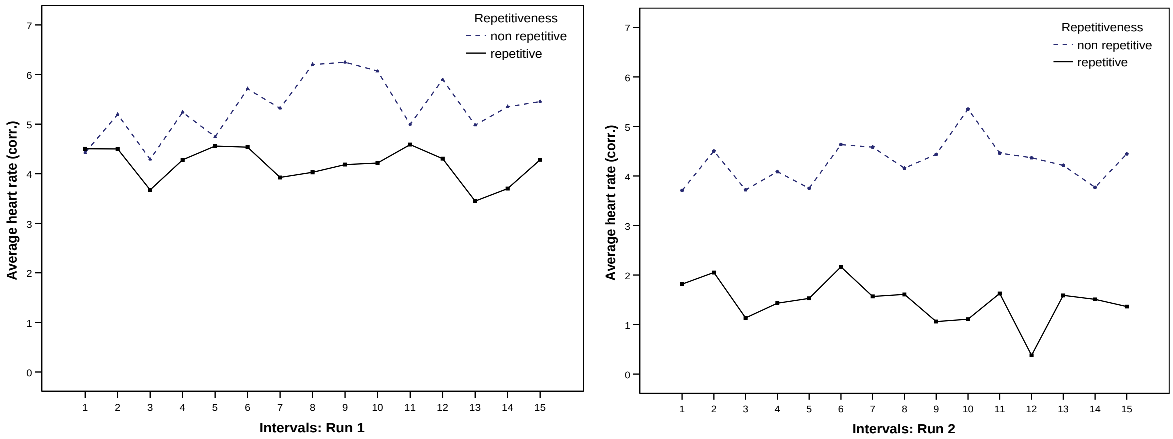


Figure 25: Average heart rate (baseline corr.) in 3-minute-intervals for each run as a function of repetitiveness (Legend: ■ black continuous line: repetitive; ▲blue dotted line: non repetitive).

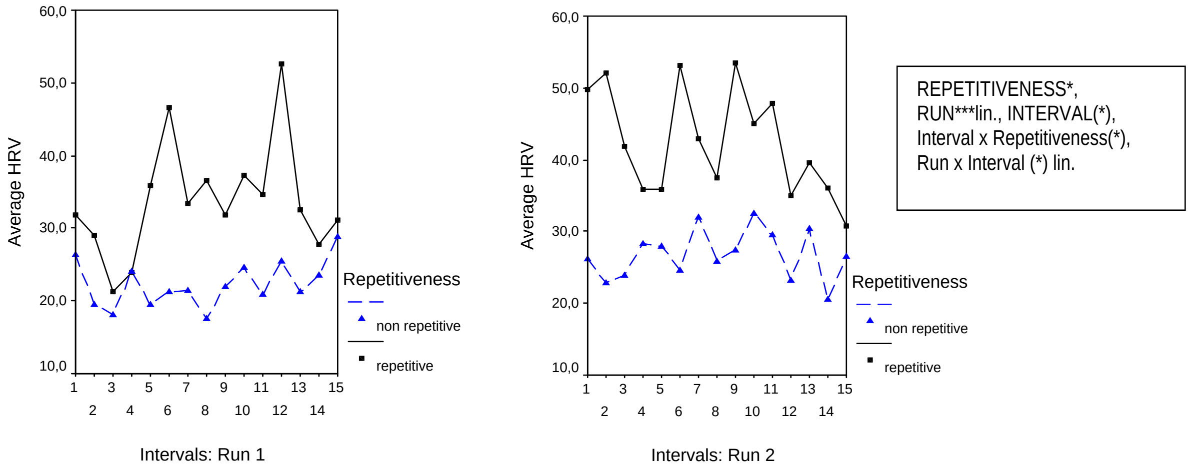


Figure 26: Average heart rate variability (in bpm) in 3-minute-intervals for each run as a function of repetitiveness.

The significant Run x Repetitiveness x DD interaction of the number of blinks (F_1=8.48, p=.010) is depicted in Figure 27 and indicates a significantly different development under low and high DD.

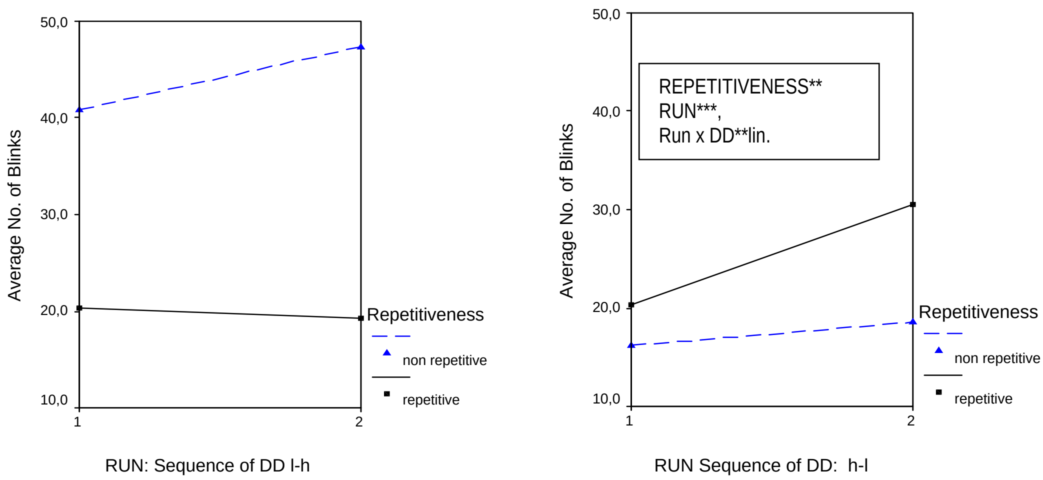


Figure 27: Average number of blinks for each run as a function of repetitiveness and sequence of dynamic density (DD): l-h: low DD in Run 1, high DD in Run 2; h-l=high DD in Run 1, low DD in Run 2.

The significant interaction in SCL indicates deactivation during the run and from the first to the second run, but no significant effect or interaction with traffic characteristics (Figure 28).

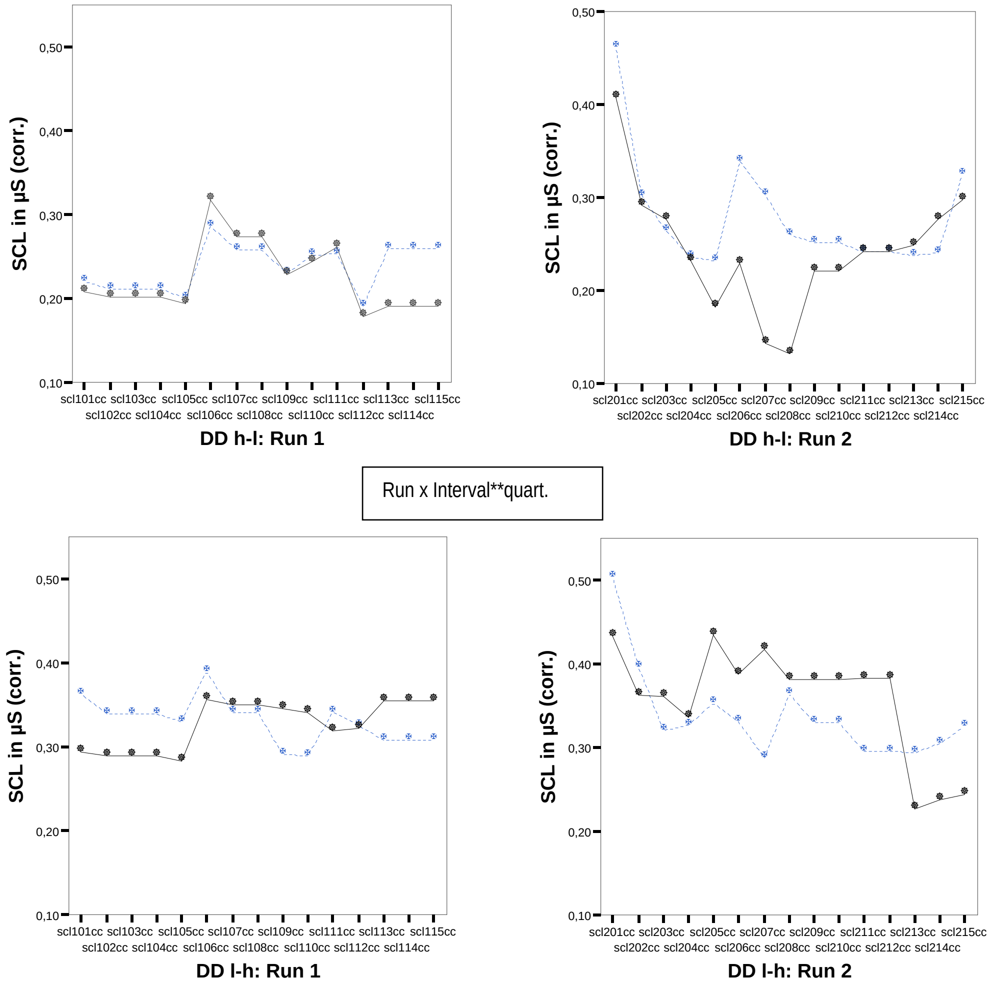


Figure 28: Average skin conductance level for each run as a function of repetitiveness and sequence of dynamic density (DD): l-h: low DD in Run 1, high DD in Run 2; h-l=high DD in Run 1, low DD in Run 2 (Legend: ■ black continuous line: repetitive; ▲blue dotted line: non repetitive).

Further analysis was conducted for the power in the EEG, wherefore the results are presented in Table 37 (descriptive statistics in Tables B-11 to B-14).

Table 37: Results of univariate Analysis of Variance for baseline-corrected EEG indicators (Study II)

Source of Variance	Results ($F_{df\ hypothesis,\ df\ error}$, p-value)			
	Theta	Alpha1	Alpha 2	Beta1
REP	*$F_{1,18}$=2.40, p=.139*	$F_{1,18}$=.17, p=.684	$F_{1,18}$=.11, p=.742	$F_{1,18}$=.24, p=.632
DD	$F_{1,19}$=.19, p=.671	$F_{1,18}$=.05, p=.834	$F_{1,18}$=.04, p=.847	$F_{1,18}$=.35, p=.559
RUN	$F_{1,18}$=4.67, p=.044*	$F_{1,18}$=7,83,p=.012*	$F_{1,18}$=.80, p=.383	$F_{1,18}$=31.62, p=.000***
INT	$F_{4.4,79}$=1.98,p=.100	$F_{2.3,41}$=2.82,p=.065	$F_{5,87}$=1.55, p=.18	$F_{2,37}$=7.91, p=.001***
Rep x DD	$F_{1,18}$=.01, p=.918	$F_{1,18}$=.00, p=.973	$F_{1,18}$=3.86, p=.065(*)	$F_{1,18}$=.64, p=.433
Run x Rep	$F_{1,18}$=.26, p=.613	$F_{1,18}$=2.70,p=.118	$F_{1,18}$=2.48, p=.133	$F_{1,18}$=.13, p=.721
Run x DD	$F_{1,18}$=03, p=.872	$F_{1,18}$=.82, p=376	$F_{1,18}$=.17, p=.685	$F_{1,18}$=.43, p=.521
Inter x Rep	$F_{4.4,79}$=1.35, p=.258	$F_{2.3,41}$=.84, p=.451	$F_{5,87}$=1.54, p=.188	$F_{2,37}$=.67, p=.520
Inter x DD	$F_{4.4,79}$=.59, p=.685	$F_{2.3,41}$=.79, p=.477	$F_{5,87}$=1.14, p=.343	$F_{2,37}$=.63, p=.543
Run x Int	$F_{5.9,106}$=1.30, p=.264	$F_{2.5,44}$=2.17,p=.116	$F_{5,90}$=2.21, p=.060	$F_{2,37}$=9.09, p=.001***
Run x Int x Rep	$F_{5.9,106}$=.84, p=.541	$F_{2.5,44}$=.91, p=.427	$F_{5,90}$=.98, p=.432	$F_{2,37}$=1.11, p=.340
Run x Int x DD	$F_{5.9,106}$=1.77, p=.113	$F_{2.5,44}$=1.29, p.=288	$F_{5,90}$=1.74, p=.133	$F_{2,37}$=.73, p=.490
Run x Rep x DD	$F_{1,18}$=.21, p=.655	$F_{1,18}$=.31, p=.584	$F_{1,18}$=.22, p=.647	$F_{1,18}$=.86, p=.365
Int x Rep x DD	$F_{4.4,79}$=1.35, p=.257	$F_{2.3,41}$=.53, p=.612	$F_{5,87}$=1.28, p=.281	$F_{2,37}$=2.16, p=.129
Run x Int x Rep x DD	$F_{5.9,106}$=.92, p=.482	$F_{2.5,44}$=.75, p=.504	$F_{5,90}$=.81, p=.542	$F_{2,37}$=1.22, p=.309

Note. N=24. Rep=Repetitiveness; DD=Sequence of DD; Int=Interval. ***p<.001. **p<.01. *p<.05. (*)p<0.1.

As shown, after baseline-correction no significant main effect was found for the between-subject factors. Significant effects in run revealed that the power in the beta band decreased from the first to the second run and increased in the alpha1 and the theta band. A tendencially significant interaction between repetitiveness and sequence of DD was found in the alpha2 band.

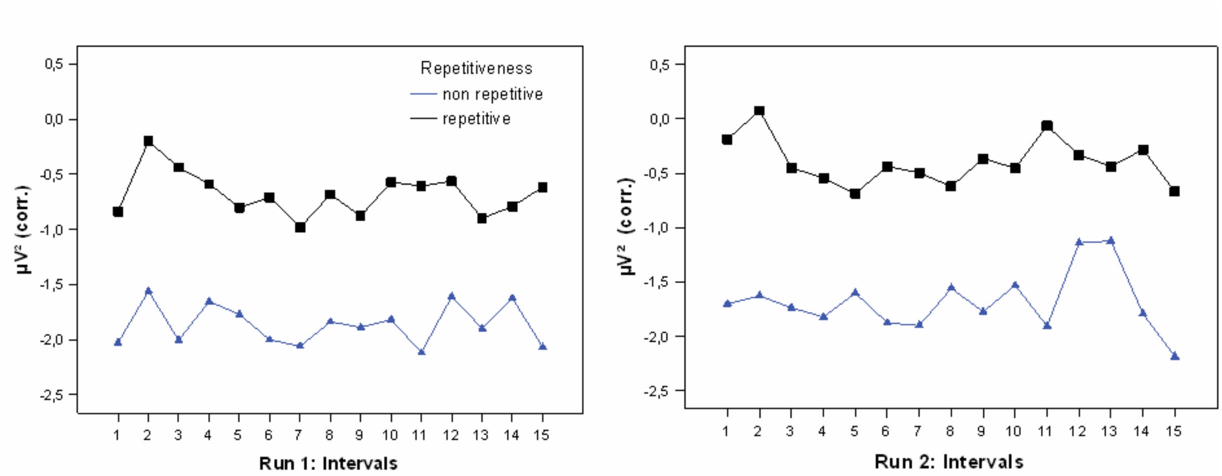


Figure 29: Average power in the theta band as a function of repetitiveness (Legend: ■ black continuous line: repetitive; ▲blue dotted line: non repetitive)

Also, the development of the power in the theta band is shown in Figure 29, as a tendencially significant factor of repetitiveness was found.

5.3.5.2. Effects of task characteristics on the subjective level

Self-reported states during scenario (TSI)

Each item of the scale was submitted to univariate analysis. The results of the univariate statistical analysis are presented in Table 38; descriptive statistics are summarized in Table B-15. Greenhouse-Geisser-corrected degrees of freedom were used in case sphericity was not provided.

Table 38: Results of univariate Analysis of Variance for physiological indicators (Study II)

Source of Variance	Results ($F_{\text{df hypothesis, df error}}$, p-value)							
	Attentiveness	Fatigue	Boredom	Irritation	Strain	Concentration	Motivation	Sleepiness
REP	$F_{1,20}=2.46$ p=.133	$F_{1,20}=5.11$ p=.035*	$F_{1,20}=.55$ p=.466	$F_{1,20}=1.10$ p=.306	$F_{1,20}=4.15$ p=.055 (*)	$F_{1,20}=3.32$ p=. 083 (*)	$F_{1,20}=5.16$ p=.034*	$F_{1,20}=10.01$ p=.005**
DD	$F_{1,20}=.86$ p=.365	$F_{1,20}=.87$ p=.361	$F_{1,20}=.55$ p=.466	$F_{1,20}=.20$ p=.657	$F_{1,20}=.57$ p=.459	$F_{1,20}=.46$ p=.507	$F_{1,20}=.54$ p=.473	$F_{1,20}=1.32$ p=.263
RUN	$F_{1,20}=5.95$ p=.024*[a]	$F_{1,20}=9.61$ p=.006** [a]	$F_{1,20}=3.25$ p=.087 (*)	$F_{1,20}=2.03$ p=.169	$F_{1,20}=.02$ p=.893	$F_{1,20}=10.89$ p=.004** [a]	$F_{1,20}=19.41$ p=.000** [a]	$F_{1,20}=14.35$ p=.001*** [a]
INT	$F_{1,40}=15.63$ p=.000*** [a]	$F_{1.4,29}=10.9$ p=.001*** [a]	$F_{1,40}=16.12$ p=.000*** [a]	$F_{1.4,29}=3.66$ p=.052(*)[a]	$F_{1.3,25.2}=.22$ p=.698	$F_{1,40}=14.94$ p=.000*** [a]	$F_{1,40}=14.78$ p=.000***	$F_{1,40}=13.52$ p=.000*** [a]
Rep x DD	$F_{1,20}=.13$ p=.725	$F_{1,20}=.99$ p=.332	$F_{1,20}=.70$ p=.413	$F_{1,20}=.00$ p=.961	$F_{1,20}=3.28$ p=.085 (*)	$F_{1,20}=.37$ p=.550	$F_{1,20}=.33$ p=.570	$F_{1,20}=.11$ p=.741
Run x Rep	$F_{1,20}=.99$ p=.332	$F_{1,20}=9.61$ p=.006** [a]	$F_{1,20}=5.14$ p=.035* [a]	$F_{1,20}=1.30$ p=.268	$F_{1,20}=.17$ p=.686	$F_{1,20}=1.53$ p=.230	$F_{1,20}=1.49$ p=.237	$F_{1,20}=.40$ p=.535
Run x DD	$F_{1,20}=2.94$ p=.102	$F_{1,20}=.03$ p=.865	$F_{1,20}=2.11$ p=.162	$F_{1,20}=.00$ p=1.000	$F_{1,20}=.02$ p=.893	$F_{1,20}=3.45$ p=.078**	$F_{1,20}=7.39$ p=.013* [a]	$F_{1,20}=8.44$ p=009.** [a]
Inter x Rep	$F_{1,40}=12.18$ p=.000*** [a]	$F_{1.4,29}=2.29$ p=.132	$F_{1,40}=13.21$ p=.000*** [a]	$F_{1.4,29}=2.27$ p=.134	$F_{1.3,25.2}=.31$ p=.632	$F_{1,40}=6.14$ p=.005** [a]	$F_{1,40}=.34$ p=.710	$F_{1,40}=.22$ p=.802
Inter x DD	$F_{1,20}=.32$ p=.727	$F_{1.4,29}=1.78$ p=.192[b]	$F_{1,40}=.57$ p=.572	$F_{1.4,29}=.17$ p=.768	$F_{1.3,25.2}=.36$ p=.603	$F_{1,40}=.08$ p=.920	$F_{1,40}=.20$ p=.822	$F_{1,40}=.07$ p=.929
Run x Inter	$F_{1,40}=.07$ p=.932	$F_{1.4,1}=.00$ p=1.000	$F_{1,40}=.04$ p=.963	$F_{1,40}=.60$ p=553	$F_{1,40}=3.60$ p=.036*	$F_{1,40}=.50$ p=.608	$F_{1.5,30}=.34$ p=.654	$F_{1.5,29}=.33$ p=.657
Run x Inter x Rep	$F_{1,40}=2.13$ p=.133	$F=_{1,1}=2.04$ p=.143	$F_{1,40}=.80$ p=.456	$F_{1,40}=1.16$ p=.325	$F_{1,40}=3.84$ p=.030*[a]	$F_{1,40}=1.02$ p=.369	$F_{1.5,30}=.05$ p=.911	$F_{1.5,29}=9.30$ p=.002**[a,b],
Run x Inter x DD	$F_{1,40}=1.14$ p=.328	$F_{1,1}=2.43$ p=.101	$F_{1,40}=1.15$ p=.327	$F_{1,40}=.42$ p=.662	$F_{1,40}=1.46$ p=.245	$F_{1,40}=.34$ p=.713	$F_{1.5,30}=1.20$ p=.303	$F_{1.5,29}=1.41$ p=.256
Run x Rep x DD	$F_{1,20}=3.60$ p=.072(*)	$F_{1,20}=.27$ p=.611	$F_{1,20}=.41$ p=.530	$F_{1,20}=.08$ p=.778	$F_{1,20}=.02$ p=.893	$F_{1,20}=3.45$ p=. 078 (*)	$F_{1,20}=5.49$ p=.030*	$F_{1,20}=.26$ p=.619
Inter x Rep x DD	$F_{1,40}=.12$ p=.889	$F_{1.4,29}=2.48$ p=.115	$F_{1,40}=.66$ p=.523	$F_{1.4,29}=.41$ p=.602	$F_{1.3,25}=1.99$ p=.169	$F_{1,40}=.47$ p=.631	$F_{1,40}=2.81$ p=. 072 (*)	$F_{1,40}=1.50$ p=.235
Run x Inter x Rep x DD	$F_{1,40}=1.89$ p=.164	$F_{1,1}=.00$ p=1.000	$F_{1,40}=.76$ p=.476	$F_{1,24}=1.16$ p=.325	$F_{1,40}=1.10$ p=.342	$F_{1,40}=1.90$ p=.163[a]	$F_{1.5,30}=.05$ p=.911	$F_{1.5,29}=4.25$ p=.035* [a]

Note. N=24 Rep=Repetitiveness; DD=Sequence of DD; Inter=Interval. Trend effects: [a]linear; [b]quartic; [c]cubic; [d]quartic; [e]order 5; [f]order 7; [g]order 8; [h]order 10; [i]order 14. ***p<.001. **p<.01. *p<.05. (*)p<0.1.

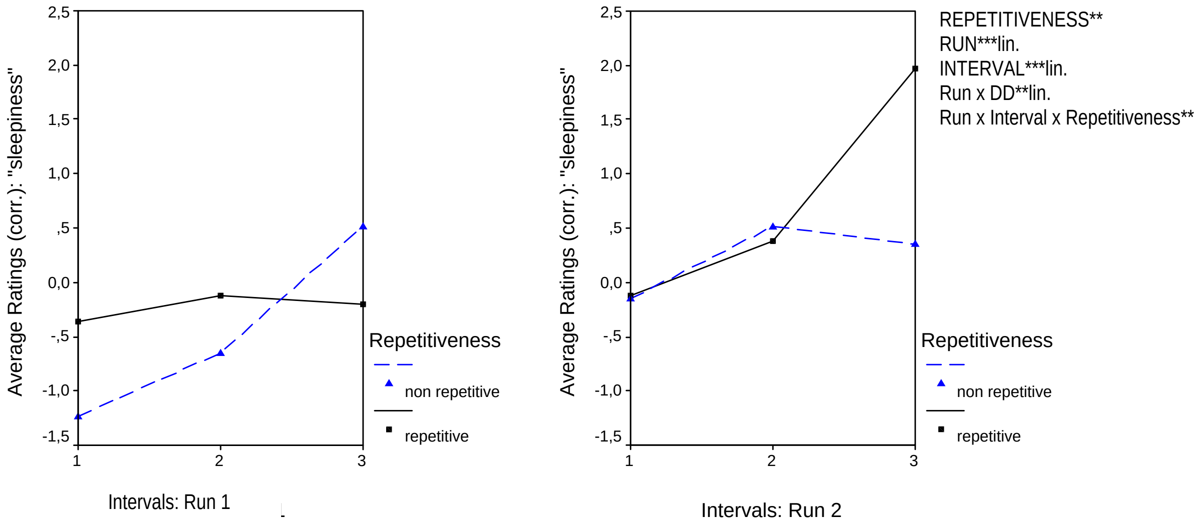


Figure 30: Average ratings for sleepiness (level corr.) in 15-minute-intervals for each run as a function of repetitiveness.

For sleepiness the significant effect of repetitiveness indicated greater scores for the repetitive group. Moreover, sleepiness increased from the first to the second run and during the scenarios. The significant interaction between run and sequence of DD can be described with a pronounced increase for the high-low-group from the first to the second run. The significant Run x Interval x Repetitiveness interaction is presented in Figure 30. Fatigue was rated higher by the repetitive group in the first run, but by the non repetitive group in the second run, which was also indicated by the linear trend in Run x Repetitiveness. The quadratic trend in the Interval x DD is due to the higher values in the low DD condition of the first run, but overall fatigue linearly increased during each run and from the first to the second run (Figure 31).

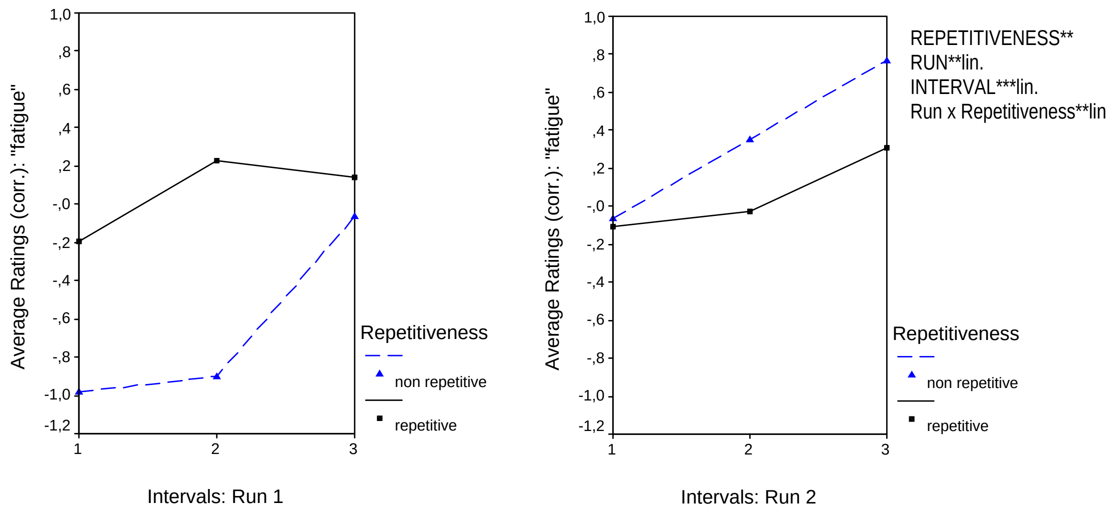


Figure 31: Average ratings for fatigue (level corr.) in 15-minute-intervals for each run as a function of repetitiveness.

Attentiveness linearly decreased from the first to the second run and within the scenarios. At the beginning of a scenario the repetitive group rated their attentiveness lower. At the end of a scenario and after a pronounced decrease the non repetitive group rated their attentiveness lower. A similar course in concentration is demonstrated in Figure 32. Whereas the main effect of repetitiveness approached significance, the significant effects of run and interval indicated a decrease of concentration within one scenario as well as from the first to the second run. Moreover, a significant interaction effect between run and interval was observed. Towards the end of the scenarios, the repetitive group rated their concentration higher.

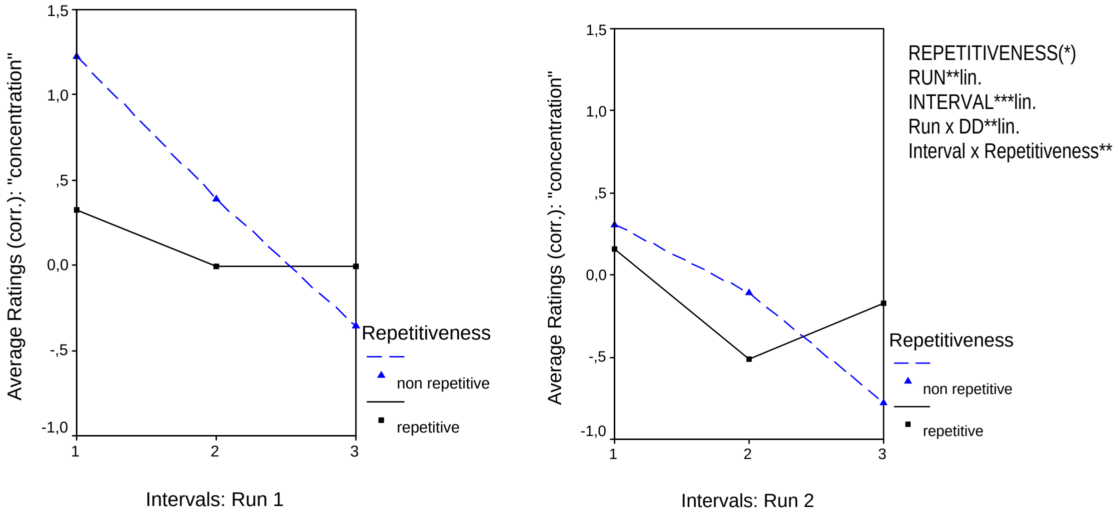


Figure 32: Average ratings for concentration (level corr.) in 15-minute-intervals for each run as a function of repetitiveness.

Boredom was increasing from the first to the second run and linearly during each run (Figure 33). The rating was higher in the first run while the non repetitive condition was higher in the second run. The linear trend component in Run x Repetitiveness indicated a different development depending on the treatment within each run, which indicated higher boredom in the first run for the repetitive condition, the linear and quadratic trend component in Interval x Repetitiveness also indicated that boredom was decreasing again towards the end in the repetitive condition, but increased in the non repetitive condition.

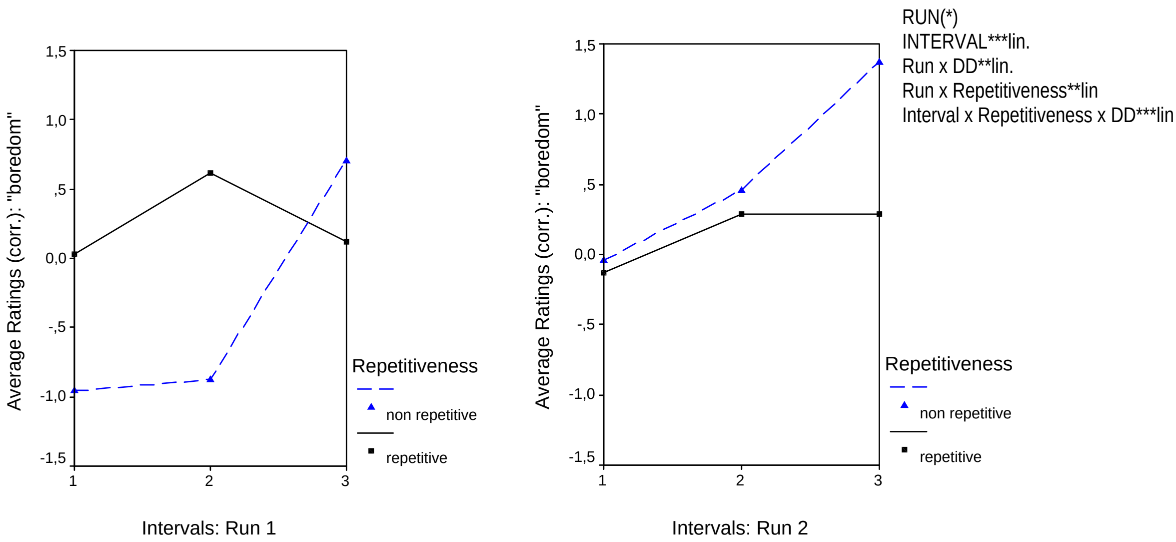


Figure 33: Average ratings for boredom (level corr.) in 15-minute-intervals for each run as a function of repetitiveness.

There was a tendencially significant main effect of repetitiveness in the ratings of strain. Strain was rated lower in the repetitive condition. The interaction effects related to the linear trend component indicated increased ratings in the repetitive condition towards the end of the first run, but decreased in the non repetitive condition, while in the second run strain increased more pronounced in the non repetitive condition (Figure 34).

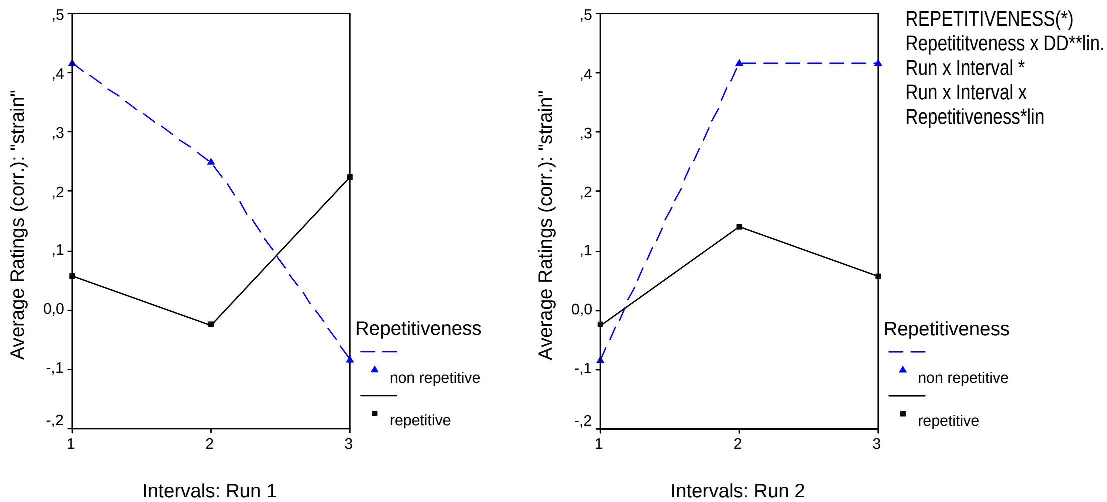


Figure 34: Average ratings for strain (level corr.) in 15-minute-intervals for each run as a function of repetitiveness.

Motivation decreased in the sequence of DD group from high-low in both repetitiveness-conditions from the first to the second run (Figure 35). The ratings of the repetitive group with the sequence low-high increased in the second run while the ratings of the non repetitive group decreased in the second run. This was supported by the significant linear trend of the DD x Run interaction.

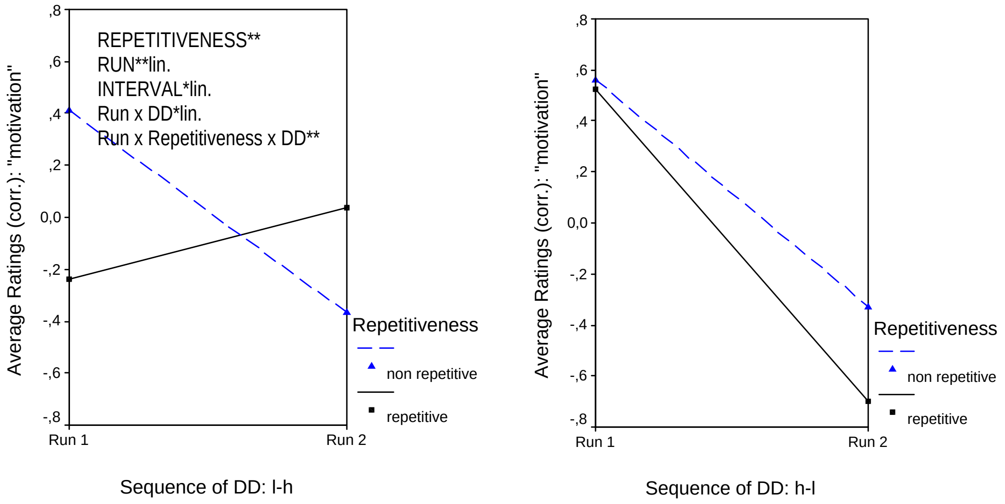


Figure 35: Average ratings for motivation (level corr.) in 15-minute-intervals for each run as a function of repetitiveness and sequence of dynamic density (DD): l-h: low DD in Run 1, high DD in Run 2; h-l=high DD in Run 1, low DD in Run 2.

Ratings for critical states after scenario (SOF)

The occurrence of critical states was assessed with a questionnaire after each scenario. Again, the results of the statistical analysis are contained in Table 39 and descriptive parameters are contained in Table B-16.

Table 39: *Results of Analysis of Variance for Scale of Feelings(SOF) subscales (Study II)*

Source	Results ($F_{df\ hypothesis,\ df\ error}$, p-value)			
	Stress	**Monotony**	**Fatigue**	**Satiation**
REP	**$F_{1,20}$=3.55, p=.074(*)**	$F_{1,20}$=.00, p=.965	$F_{1,20}$=6.22, p=.022*	$F_{1,20}$=.16, p=.698
DD	$F_{1,20}$=.01, p=.926	$F_{1,20}$=.07, p=.791	$F_{1,20}$=.39, p=.541	$F_{1,20}$=.89, p=.357
RUN	**$F_{1,20}$=4.18, p=.054(*)**	**$F_{1,20}$=13.67, p=.001** [a]**	**$F_{1,20}$=11.69, p=.003** [a]**	**$F_{1,20}$=27.59, p=.000*** [a]**
Run x Rep	$F_{1,20}$=.18, p=.678	**$F_{1,20}$=3.24, p=.087(*)**	$F_{1,20}$=.23, p=.635	$F_{1,20}$=2.74, p=.114
Run x DD	$F_{1,20}$=.71, p=.410	**$F_{1,20}$=10.39, p=.004** [a]**	$F_{1,20}$=.20, p=.663	$F_{1,20}$=2.55, p=.126
Rep x DD	$F_{1,20}$=.14, p=.710	$F_{1,20}$=.20, p=.660	$F_{1,20}$=1.07, p=.314	$F_{1,20}$=.05, p=.823
Run x Rep x DD	$F_{1,20}$=1.04, p=.319	**$F_{1,20}$=4.35, p=.050 [a]**	$F_{1,20}$=.04, p=.846	$F_{1,20}$=1.15, p=.296

*Note. N=24 Rep=Repetitiveness; DD=Sequence of DD; Inter=Interval. Trend effects: [a]linear. ***p<.001. **p<.01. *p<.05. (*)p<0.1.*

In all subscales a linear increase from the first to the second run was found. The group under the non repetitive was more fatigued than the group under the repetitive condition, as already indicated in the fatigue item. Similarly, ratings in the stress-subscale were also tendencially higher in the non repetitive group.

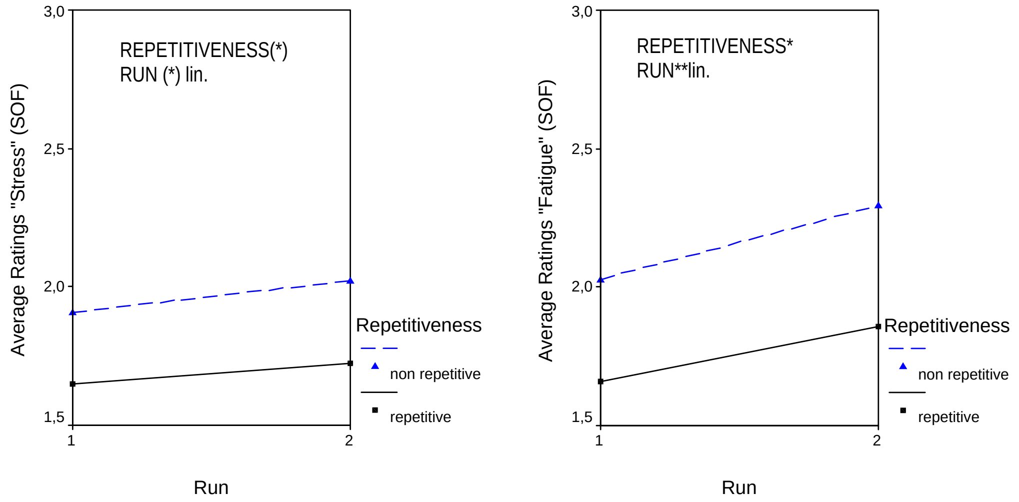


Figure 36: Average ratings for the Scale of Feelings(SOF) subscales *stress* and *fatigue* for each run as a function of repetitiveness

The ratings in the monotony subscale reflect a similar course as the individual feeling of monotony (compare 5.3.1) and are contrary to the earlier described ratings of motivation. The significant linear trend in the Run x sequence of DD interaction is a sign of the decreasing monotony from the first to the second run if the DD changed from low to high.

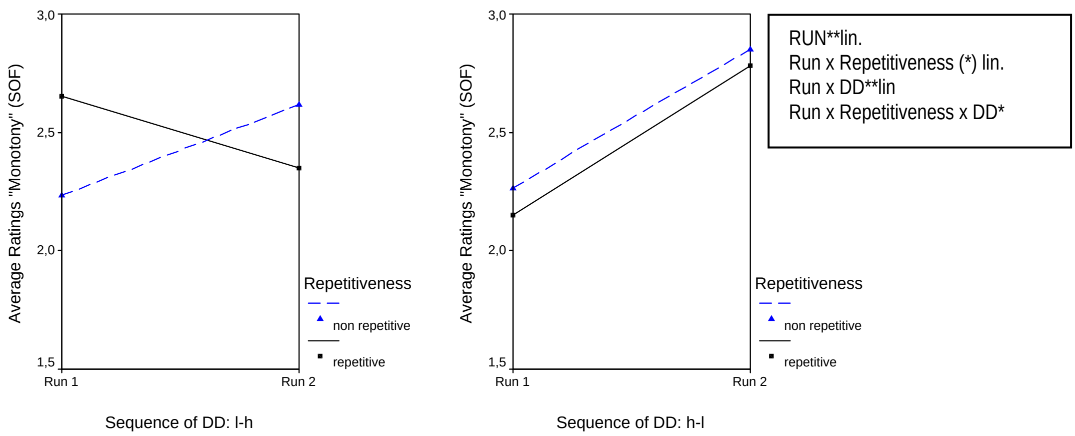


Figure 37: Average ratings for subscale monotony for each run as a function of repetitiveness and sequence of dynamic density (DD): l-h: low DD in Run 1, high DD in Run 2; h-l=high DD in Run 1, low DD in Run 2.

Ratings for mental workload after scenario (NASA-TLX)

The results of the statistical analysis for NASA-TLX are presented in Table 40 and descriptive parameters are contained in Table B-17.

The non repetitive group rated mental demand, temporal demand and effort significantly higher and subjective ratings of performance (low ratings=good performance) were lower compared to the repetitive group.

Mental demand and performance decreased from the first to the second run for the high-low-group and increased for the low-high group.

Temporal demand increased from the first to the second run for the low-high group and decreased for the high-low group. For the l-h group, temporal demand was rated higher in the non repetitive condition. No significant effects or interactions were associated with the ratings of situation awareness.

Table 40: Results of Analysis of Variance for individual items of NASA-TLX and situation awareness (Study II)

Source	Results ($F_{df\ hypothesis,\ df\ error}$, p-value)						
	Mental demand	**Temporal demand**	**Effort**	**Perform-ance**	**Frustration**	**Overall**	**SA**
REP	**$F_{1,20}$=18.67 p= .000*** **	**$F_{1,20}$=4.89 p=.039 ***	**$F_{1,20}$=15.65 p=.001*** **	**$F_{1,20}$=5.15 p=.035 ***	$F_{1,20}$=.40 p=.535	**$F_{1,20}$=5.31 p=.032***	$F_{1,20}$=2.67 p=118
DD	$F_{1,20}$=1.54 p=.229	**$F_{1,20}$=5.60 p= .028***	$F_{1,20}$=1.60 p=.221	$F_{1,20}$=.02 p=.905	$F_{1,20}$=1.46 p=.241	**$F_{1,20}$=3.09 p=.094 (*)**	$F_{1,20}$=.07 p=.801
RUN	$F_{1,20}$=.07 p=.798	$F_{1,20}$=.029 p=.866	$F_{1,20}$=.25 p=.621	$F_{1,20}$=.35 p=.559	$F_{1,20}$=3.82 p=.065 (*)	$F_{1,20}$=1.37 p=.255	$F_{1,20}$=.34 p=.566
Run x Rep	$F_{1,20}$=2.60 p=.123	$F_{1,20}$=.001 p=.978	$F_{1,20}$=1.75 p=.201	$F_{1,20}$=.14 p=.713	$F_{1,20}$=.02 p=.902	$F_{1,20}$=1.62 p=.218	$F_{1,20}$=.08 p=.786
Run x DD	**$F_{1,20}$=9.57 p=.006**[a]**	**$F_{1,20}$=7.43 p= .013*[a]**	**$F_{1,20}$=3.43 p=.079(*)[a]**	**$F_{1,20}$=5.34 p= .032*[a]**	$F_{1,20}$=1.46 p=.241	**$F_{1,20}$=5.05 p=.036 *[a]**	$F_{1,20}$=.10 p=.761
Rep x DD	$F_{1,20}$=.25 p=.621	**$F_{1,20}$=5.00 p= .037****	$F_{1,20}$=1.36 p=.257	$F_{1,20}$=.43 p=.522	$F_{1,20}$=.86 p=.364	$F_{1,20}$=.66 p=.425	$F_{1,20}$=.57 p=.460
Run x Rep x DD	$F_{1,20}$=.02 p=.898	$F_{1,20}$=.59 p=.452	$F_{1,20}$=.59 p=.452	$F_{1,20}$=.29 p=.595	$F_{1,20}$=.02 p=.902	$F_{1,20}$=.37 p=.548	$F_{1,20}$=.47 p=.500

Note. N=24 Rep=Repetitiveness; DD=Sequence of DD; Inter=Interval. Trend effects: [a]linear. ***p<.001. **p<.01. *p<.05. (*)p<0.1.

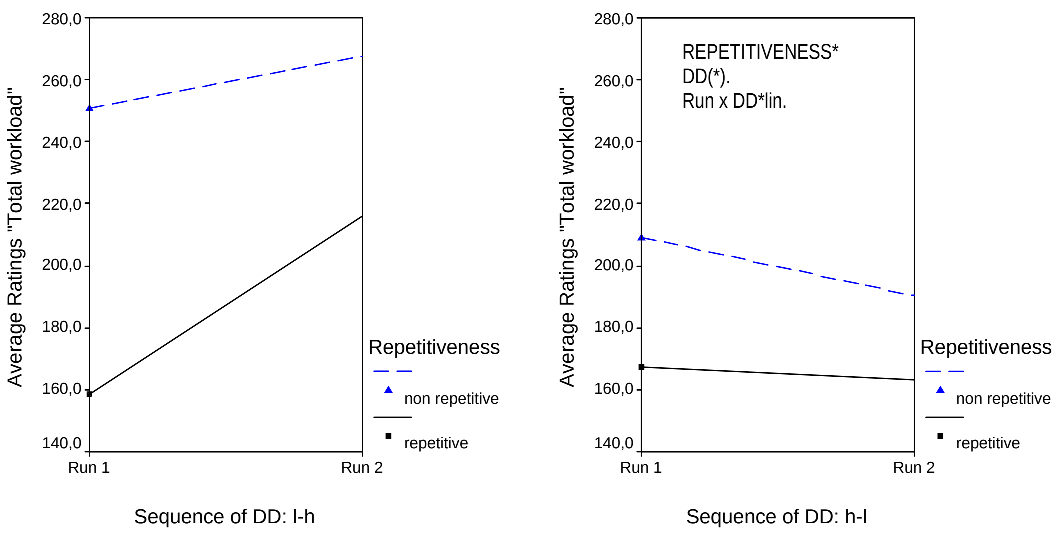


Figure 38: Average ratings for overall workload for each run as a function of repetitiveness and sequence of dynamic density (DD): l-h: low DD in Run 1, high DD in Run 2; h-l=high DD in Run 1, low DD in Run 2.

Ratings for mood after scenario (UWIST)

Subjective mood was assessed on three dimensions and descriptive statistics are presented in Table 41.

Table 41: Average ratings (and SD) for UWIST mood assessment subscales (Study II)

		Repetitiveness			
		Repetitive		Non repetitive	
	DD	l-h	h-l	l-h	h-l
HT	Run 1	2.94 (.32)	2.81 (.21)	2.94 (.32)	2.77 (.33)
	Run 2	2.79 (.19)	2.60 (.18)	2.89 (.20)	2.52 (.18)
TA	Run 1	2.92 (.19)	2.90 (.18)	2.56 (.38)	2.92 (.17)
	Run 2	2.98 (.22)	2.73 (.18)	2.69 (.30)	3.00 (.22)
EA	Run 1	1.56 (.17)	1.67 (.19)	1.73 (.48)	1.65 (.09)
	Run 2	1.63 (.24)	1.44 (.10)	1.75 (.14)	1.63 (.31)

Note. (HT=hedonic tone, TA=tense arousal, EA=energetic arousal) depending on repetitiveness (repetitive vs. non repetitive traffic) and sequence of DD (h-l vs. l-h) for n=24. Sequence of dynamic density (DD): l-h: low DD in Run 1, high DD in Run 2; h-l: high DD in Run 1, low DD in Run 2.

The results of the statistical analysis are summarized in Table 42.

Table 42: Results of Analysis of Variance for UWIST mood assessment subscales (Study II)

Source	Results ($F_{df\ hypothesis,\ df\ error}$, p-value)		
	Hedonic Tone	Energetic Arousal	Tense Arousal
Rep	$F_{1,20}$=.00, p=.950	$F_{1,20}$=2.02, p=.171	$F_{1,20}$=.91, p=.351
DD	**$F_{1,20}$=6.77, p=.017 ***	$F_{1,20}$=.82, p=.377	$F_{1,20}$=1.14, p=.299
Run	**$F_{1,20}$=6.93, p=.016*[a]**	$F_{1,20}$=.49, p=.490	$F_{1,20}$=.65, p=.430
Run x Rep	$F_{1,20}$=.06, p=.801	$F_{1,20}$=.49, p=.490	**$F_{1,20}$=5.83, p=.025*[a]**
Run x DD	$F_{1,20}$=1.22, p=.283	$F_{1,20}$=1.98, p=.175	**$F_{1,20}$=4.38, p=.049*[a]**
Rep x DD	$F_{1,20}$=.49, p=.493	$F_{1,20}$=.15, p=.703	**$F_{1,20}$=6.39, p=.020***
Run x Rep x DD	$F_{1,20}$=.35, p=.559	$F_{1,20}$=1.11, p=.304	$F_{1,20}$=2.10, p=.163

*Note. N=24 Rep=Repetitiveness; DD=Sequence of DD; Inter=Interval. Trend effects: [a]linear. ***p<.001. **p<.01. *p<.05. (*)p<0.1.*

On the subscale hedonic tone a significant main effect of sequence of DD was found. Participants rated their hedonic tone significantly higher when they executed the scenarios in the order from low to high DD, but a linear decrease from the first to the second run was found.

The subscale of tense arousal revealed significant 2-ways-interactions between repetitiveness and sequence of DD, between repetitiveness and run, and between sequence of DD and run. In the conditions of non repetitive traffic respectively low-high sequence of DD tense arousal increased from the first to the second run. Average values were generally higher for the repetitive and high-low condition, a slight decrease occurred from the first to the second run (Figure 39).

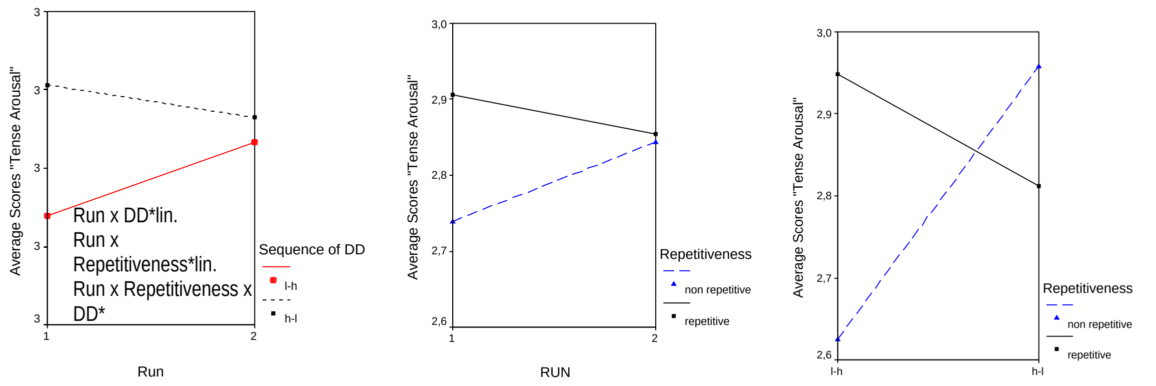


Figure 39: Average scores for the UWIST mood assessment subscale tense arousal for each run as a function of sequence of dynamic density (DD) and repetitiveness and as the interaction between repetitiveness and sequence of DD: l-h: low DD in Run 1, high DD in Run 2; h-l=high DD in Run 1, low DD in Run 2.

Also apparent is the decrease in hedonic tone from the first to the second run as indicated by its significant effect (Figure 40). No significant effects emerged on the energetic arousal subscale (all p>.171).

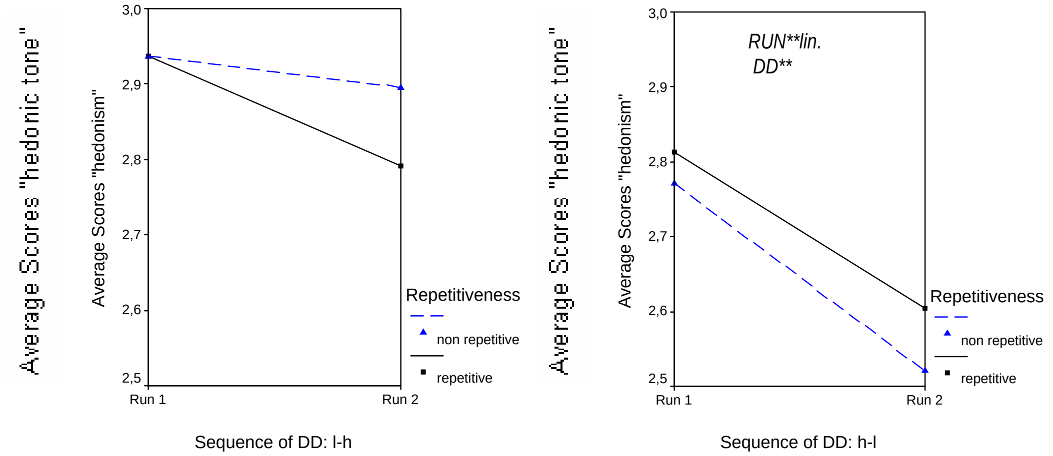


Figure 40: Average scores for the UWIST mood assessment subscale hedonic tone for each run as a function of repetitiveness sequence of DD: l-h: low DD in Run 1, high DD in Run 2; h-l=high DD in Run 1, low DD in Run 2.

5.3.5.3. *Effects of task characteristics on the behavioral level*

One of the primary performance indicators is the resolution of an out-of-routine conflict situation. To examine the role of repetitiveness in performance, the conflict resolution time and the number of STCA alerts, both related to that situation were compared in the first and second run.

Table 43: Frequency of STCA events (STCA/ No STCA) for out-of-routine conflict situation (Study II)

		Dynamic Density		
		Low	High	Total
		STCA/ No STCA	STCA/ No STCA	STCA/ No STCA
Repetitiveness	Repetitive	3/9	3/9	6 / 18
	Non repetitive	0/12	2/10	2/22
	Total	3/21	5/19	8/40

Note. (N=24, 2 scenarios)

The frequency of STCA alerts (Table 43) represented a very rare event. For this reason, the factor run was excluded from analysis and DD (low vs. high) treated as between subjects variable. The Exact Fisher Test was calculated separately (Appendix B) for each factor to examine the distributions of STCA events compared to no STCA events and resulted in no significant difference nor for repetitiveness (p=.245) neither depending on sequence of DD (p=.701).

As a second indicator the conflict resolution time from the time the conflict could be recognized until the time the first action was taken to solve it, was deducted. One subject was excluded from analysis as conflict resolution time could not be determined in one run. Table 44 displays mean and standard deviation in conflict resolution time.

Table 44: Mean conflict resolution time (and SD) for repetitive and non repetitive traffic and low(l)-high(h) vs. high(h)-low(l) sequence of DD (Study II)

Repetitiveness	Repetitive		Non repetitive	
DD	l-h	h-l	l-h	h-la
Run 1	279.83 (141.78)	287.17 (104.30)	294.83 (41.35)	153.80 (62.44)
Run 2	301.83 (86.39)	305.00 (66.40)	280.33 (110.46)	315.60 (40.13)

Note. N=23; Sequence of dynamic density (DD): l-h: low DD in Run 1, high DD in Run 2; h-l: high DD in Run 1, low DD in Run 2.[a] missing n=1.

ANOVA did not reveal significant differences for the main factors. There is a tendencially significant increase in conflict resolution time from the first to the second run (F_1=3.69, p=.070). Interactions between run and sequence of DD (F_1=3.12, p=.093) and between run, sequence of DD and repetitiveness (F_1=3.43, p=.080) are approaching significance and depicted in Figure 41.

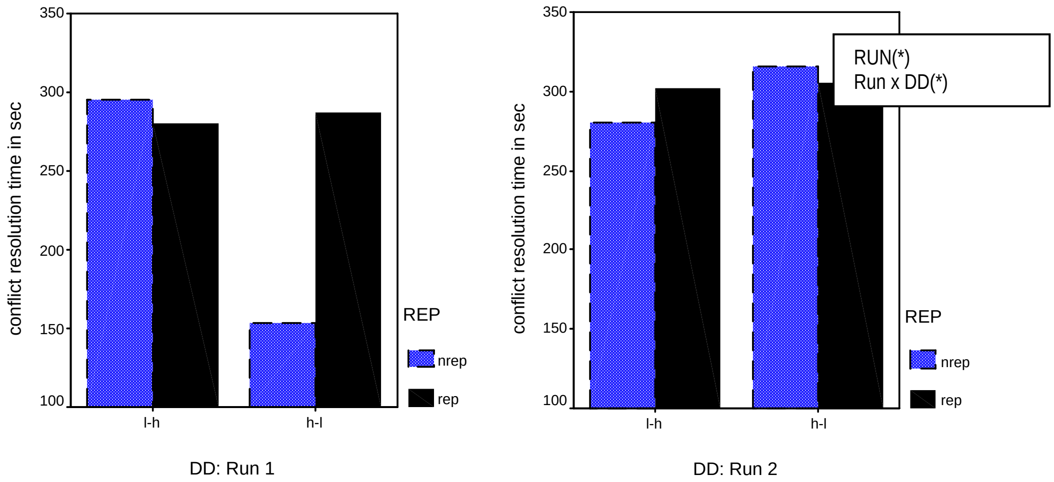


Figure 41: Conflict resolution time for each run as a function of repetitiveness and sequence of dynamic density (DD: l-h: low DD in Run 1, high DD in Run 2; h-l=high DD in Run 1, low DD in Run 2) from the first to the second run (Legend: nrep=non repetitive, rep=repetitive)

A separate univariate test was applied to the data from each run to investigate this interaction more in detail. The mean level of performance was significantly higher for the group exposed to high DD in the first run if compared to the low DD group ($F_{1,19}$=5.76, p=.027) or if the non repetitive group is compared to the repetitive group ($F_{1,19}$=5.15, p=.035).

There was no significant effect of repetitiveness on the after task performance measures assessed with the Vienna Test System. A tendencially significant effect was found in the ZBA total deviation time indicating higher deviation in the non repetitive condition (F_1=3.41, p=.085). (all p>.116; Table B-18).

5.3.5.4. The distinction of different states

The results of the development of critical states as measured with SOF were already reported in 5.3.5.2. To further evaluate the development of the states, they were compared with additional items. Based on items that were significantly correlated with critical states reported in Richter et al. (2002), appropriate items from the own study were summarized in a composed indicator for satiation and fatigue. Items combined for fatigue were the SOF items related to *exhaustion*, impaired *concentration* and *fatigue*; for satiation *irritated* from the Thackray items and *anxious* and *tense* from the mood scale were used.

A further analysis applying MANOVA investigated if these scores developed differently over time or if they can be compared with the courses of the SOF subscale indicators. For the composed fatigue-indicator, a marginally significant interaction between repetitiveness and time was found (Wilks' Lamda=.824; $F_{1,20}$=4.28, p=.052), indicating that fatigue increased in the non repetitive condition from the first run to the second run, but decreased in the repetitive condition (all results Table B-19).

5.3.5.5. The influence of additional variables

Additional variables were analyzed concerning their influence. Seventy-five percent of the participants were not of any specific circadian type, four were evening types and four were morning types. A dominance of action oriented types is indicated by 20 action oriented controllers after failure (AOF) and 22 who had high scores in decision-related action orientation; only action orientation during successful performance (AOP) was equilibrated with 12 controllers towards action and 12 towards state orientation. Concerning the personality inventory subscales the average statistics are presented in the table below (Table 45).

Table 45: Descriptive statistics for personality inventory subscales (IPIP)

	MIN	MAX	M	SD
extraversion	2.00	4.20	3.24	.69
agreeableness	2.80	4.60	3.90	.48
conscientiousness	2.40	4.70	3.57	.53
emotional stability	2.40	4.40	3.32	.58
intellect	2.80	4.60	3.65	.51

Note. N=24

The indicators of monotony were submitted to MANOVAs with factors defined by median-split in addition to the already tested task factors. The descriptive statistics and results of the statistical analysis can be found in Table B-20 and B-21. There was a significant effect of high or low scores in the intellect subscale (Wilk's Lamda=5.76, $F_{3,14}$=3.44, p=.046). A univariate analysis revealed a significant effect in sleepiness ($F_{1,16}$=5.07, p=.039). Participants who were characterized by higher open-mindedness and interest in culture scores revealed higher sleepiness scores.

Concerning the distribution of the action and state oriented performance types, the statistical analysis with the Fisher Exact test confirmed that they can be considered as equally spread on the experimental groups (p=.220). Further analysis of variance revealed no significant main effect unless the effect of repetitiveness. Also, experience as indicated by the years since fully licensed and the morningness-eveningness preference did not reveal any significant effects or interactions (descriptive statistics and results in Table B-22 and B-23).

5.4. DISCUSSION

5.4.1. The effect of Task Kactors

It was assumed that monotony as an individual state of air traffic controllers was evoked by certain traffic characteristics. The first main hypotheses investigated the effect of repetitiveness and dynamic density on selected indicators for the state of monotony based on the results in a small-scale study. The results supported the alternative hypothesis. To better represent the course over time, a standardized composed indicator is presented in Figure 42. It clearly indicates that monotony was higher in the repetitive scenarios and also higher in the low density condition of the first run.

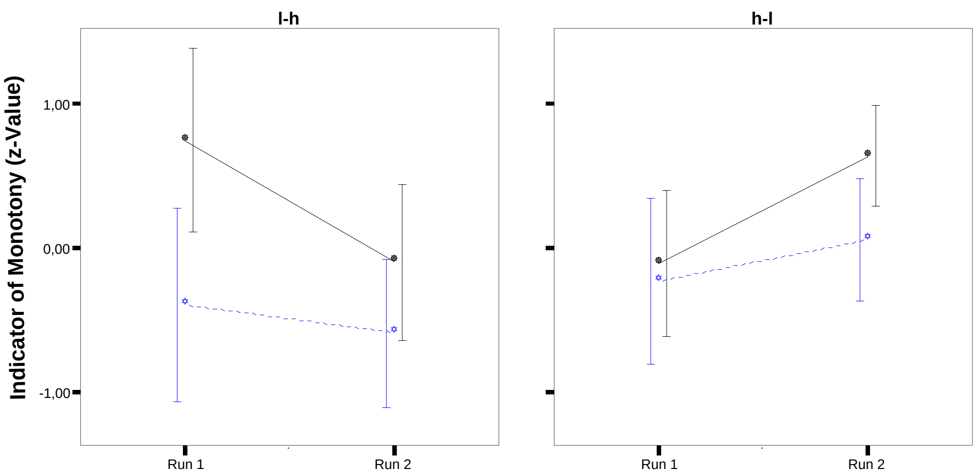


Figure 42: Average z-values and standard deviations for standardized indicator for the state of monotony as a function of repetitiveness and sequence of dynamic density (DD: l-h: low DD in Run 1, high DD in Run 2; h-l=high DD in Run 1, low DD in Run 2) for the first and the second run (Legend: ■ black continuous line: repetitive group; ▼blue dotted line: non repetitive group).

As the results were based on a multivariate analysis including HR (inv.), self-reported sleepiness and the feeling of monotony, additional analysis was undertaken for the separate indicators. Interestingly, the subjective indicators appeared to show a stronger contribution to the overall main effects compared to the physiological indicator. This indicated that the effect was perceived stronger on a subjective level but not reflected in the physiological level. Besides, the non-significant effect in HR could not be explained by individual outliers (Figure 43). However, the additional analysis of the development of the HR over time revealed the significant effect of repetitiveness. This indicated the importance of the analysis of the temporal characteristics.

A more detailed analysis of a number of variables revealed that a state of monotony could be confirmed mainly in interaction with time. On the physiological level lower heart rate and increased heart rate variability was found in the condition of repetitive traffic, no effects were found for DD. This confirmed the aspect of physiological deactivation. It is noted that the deterioration in HR appeared early in the task, as the effect of repetitiveness was already visible in the second half of the first run. While in the small-scale study an effect of DD on HRV was found, this result was not replicated in the main study. There was however a significant effect of repetitiveness. This might have been a consequence of the different levels for the DD factor in both experiments.

On the subjective level, decreased cardiovascular indicators were linked with reduced ratings of workload in the NASA-TLX subscales. Even though these findings might suggest a workload reducing effect of repetitive traffic, a different picture appears if additional indicators are considered. Also, this finding cannot be justified by lower requirements and lower traffic load in the repetitive condition, as these arguments were taken into consideration during the scenario design. Throughout the runs it was continuously necessary to monitor, control, and implement conflict resolutions in the sector. Further impacts were related to decreased concentration and increased fatigue in repetitive traffic. While controllers felt already more fatigued and bored in the repetitive scenario during the first run, this difference disappeared during the second run because of a general increase of fatigue with time-on-task. An inverted pattern is reflected in concentration. Unexpected results were found in the mood subscales and the subscale assessing the critical state of monotony. The decrease in hedonic tone and the increased tense arousal might be related to increased satiation. Even though descriptive values indicated decreases in the repetitive conditions of the subscale *energetic arousal*, its insignificant result might have been influenced by manipulations in DD.

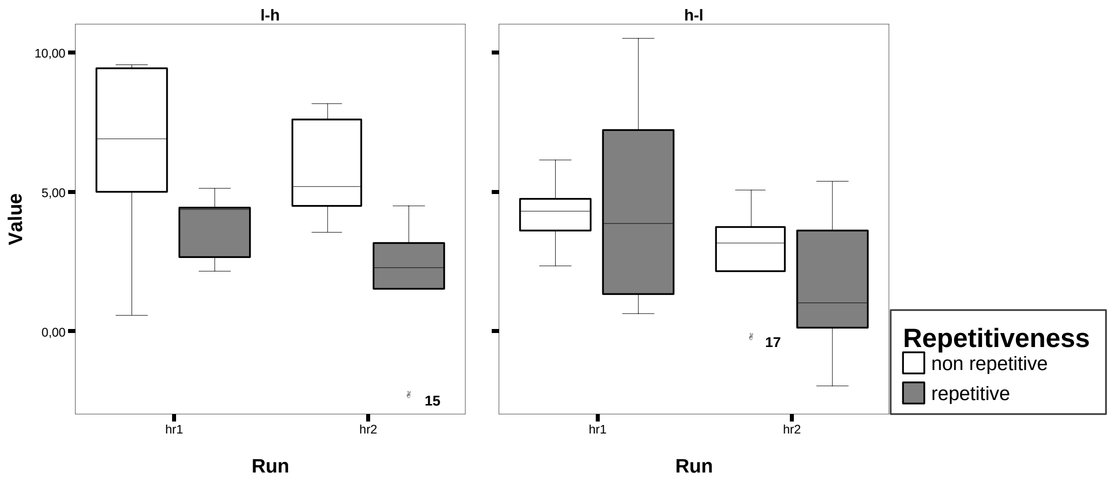
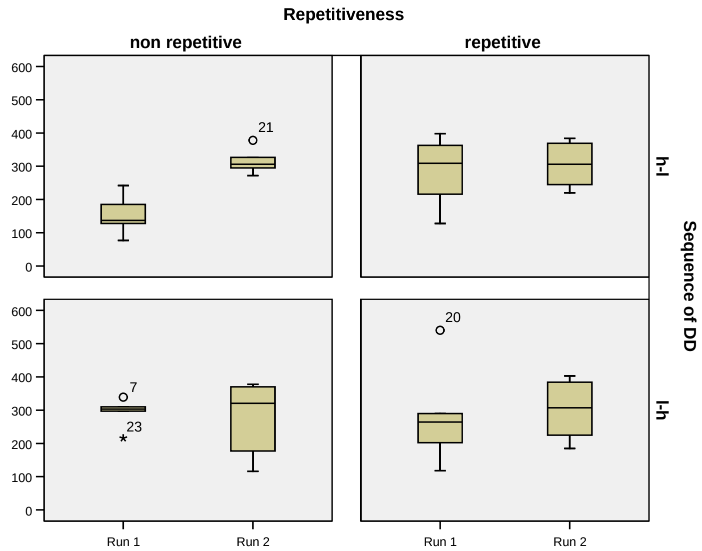


Figure 43: Box plots for corrected heart rate as a function of repetitiveness and sequence of dynamic density (DD: l-h: low DD in Run 1, high DD in Run 2; h-l=high DD in Run 1, low DD in Run 2) for each run (higher values indicate higher heart rate).

Another interesting finding is related to the subjectively perceived performance. Controllers felt that they performed better in the non repetitive condition and in the situation when the dynamic density was high in the second run after being low in the first run. In addition, the controllers felt more motivated in the non repetitive settings. The change in DD from high to low did further decrease the motivation, while the increase from low to high DD contributed to higher motivation. The results point to the aspect that controllers are challenged by the aspect of solving problems. Even if objective performance was the same, controllers only have the feeling to perform well when they had dealt with a certain amount of potential conflicts. This is supported by anecdotic descriptions of controllers, but not considered in the development of ATC concepts.

Consequently, this brings up the question how the subjective results relate to objectively assessed performance. The multidimensional assessment of a state of monotony as proposed by Bartenwerfer also predicts impaired performance**.** For this reason, the conflict resolution time and frequency of STCA events in an out-of-routine conflict situation were studied and revealed that conflict resolution time increased from the first to the second run and was tendencially higher in repetitive conditions. Low mean values found in the group that performed the first run in non repetitive high DD conditions are not caused by individual outliers (Figure 44). Remarkable is also the significant increase of the conflict resolution time in the non repetitive condition changing from high to low from the first to the second run. Furthermore, the values in the repetitive traffic condition generally demonstrate a wider range.



A REVOIR - A REVOIR – REVOIR – A REVOIR

Figure 44: Box plots for the conflict resolution time for each run as a function of repetitiveness and sequence of dynamic density (DD: l-h: low DD in Run 1, high DD in Run 2; h-l=high DD in Run 1, low DD in Run 2).

The distribution of STCA alerts completes this picture. The higher number of STCA alerts in the repetitive conditions was not significant. A significant effect might have been found if experimental conditions would have been more demanding. Failures which require recovery even occur if controllers are very well trained for dealing with unexpected situations, as Thackray and Touchstone (1983) noticed. Various factors on an organizational level or regarding individual preconditions are influencing the capability to deal with unexpected situations or dynamic changes in the traffic demands.

Because of the delayed task-effects of performance, additional performance tests were included in the set-up. There was no effect on reaction time, estimation abilities and concentration in the tests. This means that controllers perform well even after loading conditions independent on their nature.

According to the findings reported by Hockey (2003) or Mulder et al. (2003), it might be argued that the performance was maintained under increased effort, and thus would result in increased physiological costs. To investigate this hypothesis, further analysis was undertaken for the HR and HRV indicators during the performance tests (reported in Table B-24). In relation to the baseline-measures the physiological measures show a decrease of HRV in the repetitive traffic, which might indicate higher effort after repetitive traffic. Moreover it is interesting that in ZBA HR increased in repetitive and non active conditions. The high level of HR under the active condition might be a consequence of activity in rest break. This confirms again that controllers can perform very well after exposure to repetitive conditions, even though at high physiological costs. This supports the assumption of a potential relationship between short- and long-term effects of task repetitiveness. Therefore, there is a need to work on models that explain consequences of repetitive and uneventful work situations on controllers on a short-term and long-term basis in consideration of loops for mitigation effects. The question remains whether the effect on after task performance would have been found in tasks of longer duration or if different cognitive abilities would have been assessed.

As far as the physiological and subjective indicators were concerned, the main study replicated the results of the preliminary study. Both experiments showed lower heart rate in repetitive traffic. In contrast to the small-scale study, no effects of DD were found in HRV. This questions if the DD manipulation was sufficiently pronounced in this experiment. Because of a different sample of active controllers who were used to deal with a high amount of traffic, a greater difference between the conditions might have been needed. But still, an effect of the sequence was shown in other indicators. The question is whether the results can be explained by different work strategies when varying task demands, but there was no indication stated in debriefings. The differences in physiological reaction patterns from the prestudy to the main experiment might also be explained by different coping styles in terms of Frankenhaeuser's (1986) active/passive and with/out distress dimensions. Active controllers might be used to different styles in reacting to underloading situations. As was observed and reported in the debriefings, some ATCOs used strategies to remain active. One participant re-sorted flight labels on the screen even though he had the feeling to have everything under control.

To a certain extent, the results can be compared to those of Thackray et al. (1975) who investigated physiological, subjective, and performance changes accompanying reported boredom and monotony in a complex visual monitoring task. They found a similar cardiovascular pattern for the group with high ratings in feeling of monotony and boredom. Participants who reported high monotony and boredom also showed greater increases in response times, HRV, and strain while attentiveness decreased. Also, they rated their attentiveness lower and showed performance impairments. While their interpretation focused on reduced attention, the offered explanation of the results of the own study rather supports a general impairment in the individual state on multiple levels. That participants performed better in an unpredicted task after active monitoring was reported by Johansson et al. (1996). This is consistent with the finding of faster conflict resolution in the non-repetitive high density scenario. An advantage of the present study is that the sample consists of air traffic controllers and the simulation environment offered a better representation of reality compared to the environment used in the Thackray study.

These results do differ in some respects from the ones summarized in Scerbo (2001) who found also an impact on irritation and increased strain. Previous results using vigilance tasks were not confirmed in the own study, as far as the state of monotony was not only related to the decrease in attentiveness but also to decreased workload. Even though, in the own study the boredom ratings differed. This might be explained by individual differences to perceive something as boring or by the selected sample in the study of Scerbo, which cannot be compared to ATCOs. Schroeder, Touchstone, Stern, Stoliarov and Thackray (1994) found that boredom increased not only from the beginning until the end of a 2-hr-session but also across days. Hitchcock, Dember, Warm, Moroney, and See (1999) noted the effect of boredom on monitoring performance.

It is noted that the data also contain a certain pattern of association and dissociation between the individual items and the physiological indicators.

To summarize, the results support the theory of monotony as it was presented in the literature also for ATC. Under the same preconditions as introduced by Bartenwerfer, namely that the task was perceived as easy, effects could be demonstrated on various dimensions. Whereas the physiological deactivation pattern is rather clear, subjective responses are more complex. The sequence of DD was an important mediator for motivation and hedonic tone, as it developed favorably in the condition from low to high DD. This turned out to be an important factor, as favorable effects were also seen in other indicators. Therefore, it cannot be neglected within the discussion of potential countermeasures.

5.4.2. The influence of Individual Factors

The second group of hypotheses considered the influence of individual factors on monotony. Through the consideration as a further variable the significant differences depending on high or low boredom proneness or high or low initial state of recovery were found and the null hypotheses were retained.

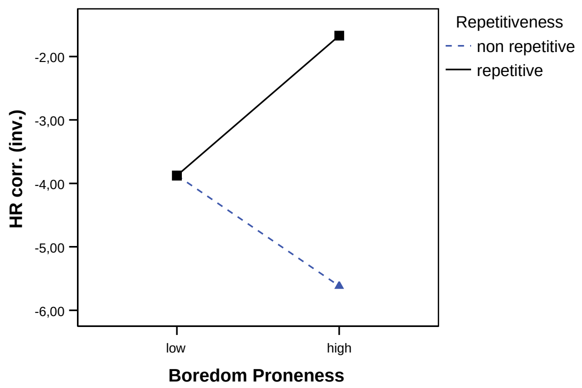


Figure 45: The development of the inverted corr. heart rate as a function of boredom proneness for the first and second run (Legend: ■ black continuous line: repetitive; ▲blue dotted line: non repetitive).

Considering the univariate analysis of the single indicators for the state of monotony, there was however a significant interaction between boredom proneness and repetitiveness in the inverted heart rate measure ($F_{1,16}$= 4.67; p=.046) and a tendencially significant interaction between boredom proneness and run in the sleepiness indicator ($F_{1,16}$=3.14; p=.096) which indicated that this subject required further investigation. Further analysis of these results indicated that controllers high in boredom proneness showed lower activation in heart rates and reduced scores in sleepiness (Figure 45). That this difference was not reflected during the second run is a further indicator for the impact of time-on-task.

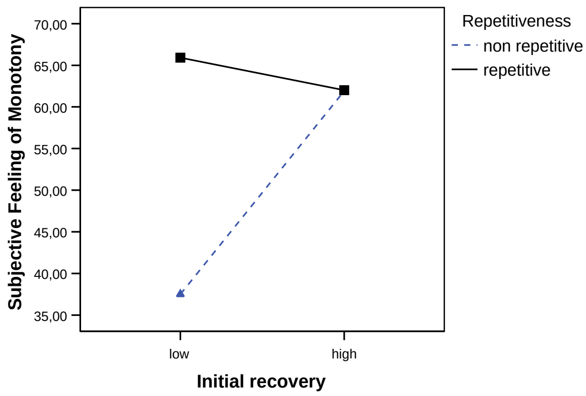


Figure 46: The development of the subjective feeling of monotony as a function of initial recovery state and repetitiveness (Legend: ■ black continuous line: repetitive; ▲ blue dotted line: non repetitive).

Also the impact of the initial state of recovery requires further investigation, as the interaction effects. For example, univariate analysis revealed a significant interaction between the initial state of recovery and repetitiveness in the subjective feeling of monotony ($F_{1,16}$=4.47; p=.050). Under the condition of lower initial recovery, participants in the repetitive condition rated the feeling of monotony higher.

5.4.3. The Impact of Countermeasures

The results for the impact of break activity on monotony supported the alternative hypothesis. The indicator for the state of monotony was lower in a final scenario, if physical exercises have been executed in a short rest break.

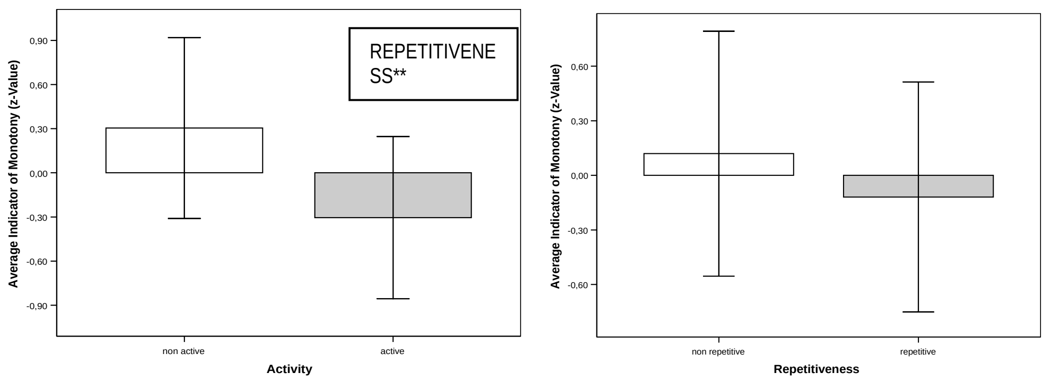


Figure 47: Average values and standard deviation of the composed indicator for the state of monotony measured in Run 3 as a function of break activity and repetitiveness.

As presented in Figure 47, the composed score of the monotony indicators was lower after the active break activity, but the participants still had higher scores in monotony after repetitive traffic. The effect is caused by the effect of sleepiness and HR. The participants, who executed active breaks, felt less monotony during the execution of Run 3.

Still, it might be the case that activity rather acted against fatigue than against monotony as indicated by the significant effect of repetitiveness. The group exposed to repetitive scenarios had a higher indicator of monotony despite the activating impact of breaks. According to the reported importance of active breaks, it might have been expected that a stronger effect would have been seen after the repetitive task. There was no indication that activity had a reinforcing effect especially after the repetitive traffic leading to a state monotony. On the contrary, monotony remained high even after the active break. This cannot be explained by influences of the sequence of DD, but by the fact that the activity had been introduced rather late. To better understand the effect, additional analysis of variance was conducted for the individual cognitive and motivational aspects. The results summarized in Table B-25 showed that the activity also favorably affected self-reported attentiveness, concentration, sleepiness, and motivation. Also, exercises were not included in form of frequent short rest breaks, which might have counteracted the monotony development already from the beginning. In the current set-up it would have disturbed the manipulation of the main experimental conditions.

Moreover, it was not possible to investigate how long the activity effect lasted in the third scenario, as the final part of the session was too short for that purpose. It can be assumed that it still affected the physiological indicators in the performance tests. Thayer (1987) reported that the effects of a brisk walk can last up until two hours. A delayed response was found in the study of Straussberger and Kallus (2003), indicating an activating effect after the end of the task. Also, there might be a difference between light physical activity and extensive sport in longer breaks. Oweis and Spinks (2001) reported unfavorable effects of intense physical activity, which might also be the reason that Kozena et al. (1996) did not find the effect of physical activity. Certainly, physical activity is a well-known mitigation strategy by many air traffic controllers, however, empirical support has not been offered yet, neither for a systematic introduction in rest breaks nor for short exercises during working at position.

5.4.4. One Step Closer – Related or Independent Critical States?

One further focus of interest was the question which additional characteristics help to distinguish similar critical states. As it was exposed, arguments support distinct states of fatigue and satiation in the context of monotony. In the following section it is discussed if the data support this assumption. For an overview, Table 46 summarizes the development of the different indicators and is used for the understanding of the critical states based on the results in the factor repetitiveness.

Table 46: Summary of the effect of repetitiveness in measured indicators

	RUN	INTERVAL	RUN 1	RUN 2
Heart rate	↓		repetitive ↓	
HRV	↑		repetitive ↑	
Number of blinks	↑		repetitive ↓	repetitive ↑
SCL	↓	↓		
EEG: theta	↑			
EEG: alpha1	↑			
EEG: alpha2			rep↓	
EEG: beta1	↓			
Subjective feeling of monotony [a]	↑		repetitive ↑	
Subjective sleepiness	↑		repetitive (i2,i3)↑	repetitive (i3)↑
Subjective fatigue	↑		repetitive ↑	
Subjective concentration	↓		repetitive ↓	
Subjective attentiveness	↓			
Subjective irritation				
Subjective boredom	↑	↑	repetitive ↑	
Subjective motivation [a]	↓		repetitive ↓	
Subjective strain			repetitive ↓, repetitive (i3)↑	repetitive (i2,i3)↓

	RUN	INTERVAL	RUN 1	RUN 2
Overall workload (TLX) [a]			repetitive ↓	
Subjective performance (TLX) [a]			repetitive ↓	
Mood: tense Arousal [a]			repetitive ↑	
Mood: hedonic tone	↓			
SOF: satiation	↑			
SOF: fatigue	↑		repetitive ↓	
SOF: monotony [a]	↑			
SOF: stress	↑		repetitive ↓	
Conflict Resolution time [a]			non repetitive ↓	

Note. [a]complex interactions with dynamic density (DD). Increasing (↑) and decreasing (↓) changes in experimental condition; i=interval; TLX= Task Load Index; SOF=Scale of Feeling.

That overall workload was tendencially rated lower and the subjective feeling of monotony was significantly higher in the repetitive condition is consistent with the theory of Bartenwerfer, which predicted that an easy task provokes monotony.

A different development over time may also support a distinction of critical states. As it was mentioned in the literature, satiation and monotony can result immediately already at the beginning of a task, while fatigue is associated with emptying resources during task execution. Also, the data indicate that monotony develops rather soon, whereas after a longer time-on-task general fatigue overlaps with the effects of repetitiveness. During the first simulation run of the experiment repetitiveness had an impairing effect on perceived fatigue, concentration and sleepiness; in contrast fatigue dominates during the second run, where the values continuously increased or decreased and between-group differences disappeared. Another support for the different states is the different effect of DD as reflected in the SOF subscale of monotony, where an increase of DD led to less monotony, while at the same time fatigue was increasing. In a call center, differences in monotony and satiation at the end of the working week also indicated the cumulative effect of exposure to work (Richter et al., 2002).

Increased effort was not sufficient to explain this result, as effort remained low. Thus, it was hypothesized that the differences or similar reaction patterns in differently described states would indicate a better base for the distinction of critical states than the historically distinguished different states included in the International Standards for Mental Workload (ISO 10075). Two approaches were followed. Due to the low number of participants results of correlations may only be used for the purpose of hypotheses-generation and are reported in Appendix B.2.7.

First, the indicators which were correlated with the critical states in Richter et al. (2002) were combined to indicators for critical states. The problem of the SOF was that the subscales are not independent from each other, while the composed indicators did not correlate with each other. Still, they were related with SOF subscales, except for the satiation indicator. This was surprising, as theoretically the items represented the concept. The second approach analyzed the intercorrelations of all assessed items for the first and second run (Table B-27) and systematized similar groups. This was also one of the procedures proposed by Leonova (2003), and helped to better present a basis for a theoretical framework, which does not mix up all the related concepts around monotony, but clearly assigns certain terms and notions to a certain phenomena around a state. This is especially relevant for terms which seem to represent similar aspects.

Also, the development of items representing satiation need to be compared with items around these concepts such as boredom and motivation. None of the satiation indicators shows similar patterns compared to tense arousal or hedonic tone. One reason might be that boredom seems to be a process that appears in monotony and in satiation due to a lack of alternative interpretations. An operator may state to be tired to express low motivation to continue a task. This is in agreement with arguments from Briner (1999) that people mix up their emotions. A similar process might take place in attentiveness and concentration, which are both fatigue-related but not reverted to the judgment of sleepiness.

After-effects measured with performance tests were planned to distinguish between fatigue and monotony. However, as it was not possible to introduce a pretest, a clear support for the fatigue concept cannot be deducted from the post-tests. However, the discussion of states of monotony and fatigue as independent concepts should consider the different effects of the break activity, which did not counteract the monotony in the repetitive condition.

Finally, the debate around critical states has not addressed that the investigated task characteristics might also be related to positive states. The demand of many controllers to experience challenge supports the consideration of the concept of flow, which helps to understand how controllers cope with their job. As such, this aspect is also neglected in concept development, where the focus is set on performance optimization, but not on the facets that contribute to avoid performance breakdown. A consideration of positive states might thus help to understand which mechanisms underlie monotony resistance. The experience of flow might also have counteracted the development of monotony in the current study. For this reason, the impact of a state of flow was included in the interpretation. Close to the flow concept of Czikszetimihalyi (1975), the major aspects were combined in an indicator of flow experience, that is hedonic tone (happy), energetic arousal (active), concentration (cognitive efficiency) and motivation. A median-split was used to create high and low flow experience and submitted as a further factor in the analysis of variance. The results (reported in Table B-26) revealed that flow experience in fact counteracted the occurrence of monotony. No effect of monotony was found if participants had a higher value in the flow indicator. This was relevant in the first run, but not in the second run. Thus, this aspect deserves further investigation.

5.4.5. Methodological Issues

When interpreting the results, several methodological issues should be considered, as they do illustrate some limitations of the present study. First of all, the applied questionnaires and scales were useful to gain a broad picture of the reaction pattern. The first measurement point in the subjective ratings was not included in the analysis for a better comparison with the results of the small-scale study. This did however not have a major influence on the results. Decreasing the scenario from 50 to 45 minutes did not eliminate the effect on any of the collected measures.

The ratings on the monotony subscale reflect a similar development as the individual feeling of monotony. This supports the assumption that a few indicators are sufficient to detect monotony states. The questionnaire for the assessment of critical states was not reflecting the expected results. This was already elaborated upon in the discussion of the small-scale study results. It might be a consequence that this questionnaire was applied in a simulation environment, even though developed for field settings. Despite the development of a trait and a state version, the application of the state version was not sufficient to reduce the artificial set-up effects. Richter et al. (2002) stressed that SOF only is useful when work periods are longer than 4 hours. It also could be criticized that no baseline was collected for the mood and critical states. Not being applicable for the SOF in simulation settings as it assesses task execution, it might have been considered for the mood assessment. The results in the SOF subscales are similar to the intercorrelations of the subscales reported by Rockstuhl (2002), which is interesting for further hypothesis-generating considerations.

This results in the discussion if the critical states assessed with SOF can be seen as independent, or just reflect that a different focus is set on emotional, cognitive and motivational levels. A clear indication on how these states interrelate helps in their assessment and the selection of countermeasures. As already criticized, SOF contains task and state-related items, which are not clearly distinguished. It would be better to describe either the task or someone’s feelings, but not to confuse both in one scale. This is also valid for the boredom rating, as it just asked for an overall boredom rating, but not how boring the task, or the scenarios, or the environment were. This could explain unclear results in boredom, as context might have strongly affected the rating, and which facets were focused in the moment of the assessment.

The current investigation was conducted with a group of active ATCOs. Even though, the task execution during the experiment was new and thus a change against everyday work, which seems to suggest that the effects of monotony were assessed rather conservatively. This was especially the case with one participant who said that he actually woke up during the task execution. Thus the results may still underestimate the potential of monotony in such conditions. But statements collected in the debriefing session supported the perception of the traffic as repetitive and monotonous. A final question concerns the generalizability from the sample. The participating center is of a very special nature because of the variety of nationalities in the employed controllers and of the generally high traffic load as a European crossing point. Especially if the results are compared with results from earlier studies, the multinational composition might have had an impact on the results. This might affect the fine-tuned understanding of some of the items in the scales. Even though the participants were partly of English mother tongue and otherwise had a good working level of English, some of the words might have a slightly different meaning depending on the original cultural background. For example, this was already discussed for the terms around stress and strain in the literature background.

Overall, the method appeared appropriate to investigate the defined task and individual factors in the simulation environment concerning their potential to evoke monotony. In a next step this will be investigated in the operational field.

5.4.6. Recommendations for Next Steps

Even though, some of the research questions are already answered in the simulation set-up, several factors will be considered in a field approach. This is necessary to investigate if the found task factors do show comparable results in the operational environment, but also a variety of additional factors have not sufficiently been considered yet.

One of them is the perceived predictability of the traffic. This is related to the assumption that if something is already predicted as more repetitive or uneventful at the beginning it would reinforce the experience of monotony and thus contain an implicit safety-relevant error potential. This might be similar to the outcomes in a study reported in Rau and Richter (1996, p. 275), where the anticipated demands had a greater influence on psychophysiological strain than the subsequent evaluation.

A related aspect are habits, which indicate the way of dealing with routine and also how do deal with routine if extraordinary conditions occur that require to deviate from habits which have been formed through experience. The role of routine will be specifically considered in the next study.

The previously deployed materials will be integrated again in the study. Despite the shortcomings discussed for SOF it will be used to investigate if the same results will be found in the operational field. Because of the complex cardiovascular system, where arterial BP is the most important factor as it is regulated by cardiac output and afterload, the assessment of blood pressure might be considered, as HR is also just one of the important variables in the complex cardiovascular system.

Finally, a further focus is placed on the occurrence of positive states related to the control task and which mental sets support the maintenance of optimal individual states for the task. To encourage positive states through appropriate initiatives has not been considered as a resource yet because of the tendency in psychology to rather focus on problem situations than on positive aspects related to work, which only gained relevance within the German work psychology.

5.5. SUMMARY

The scope of the simulated study was to determine factors that evoke and influence monotony in ATC. It was hypothesized that the repetitive activity would have an impact on multiple indicators on the physiological, subjective and behavioral level, wherefrom the most effective ones were composed into standardized indicators for a state of monotony after the small-scale study. Furthermore, it was expected that the low or high sequence of dynamic traffic density would affect monotony. The design of rest breaks was assumed to be a potential countermeasure that is easy to implement in the operational environment. Individual factors might however interfere with the effects on task characteristics.

An experiment deploying a 2 x 2 x 2 x 3 (vs.15)-mixed design with repeated measures on the last two factors was conducted. Twenty-four controllers executed two traffic scenarios of 45 minutes each and a short third scenario to determine the effects of break activity. The dependent variables comprised heart rate and its variability, skin conductance level, and blink activity. On a subjective level, scales assessed mood, workload and the perceived cognitive, emotional, and motivational state during and after the scenarios. The results support the main hypothesis for the task effects. The indicators for a state of monotony were higher if participants were exposed to repetitive scenarios. The effect on monotony was consolidated in the low density condition of the first run.

The comparison of indicators for the critical states revealed that a state of monotony as a consequence of task repetitiveness was clearly found in the first scenario, but overlaid by time-on-task effects resulting in higher fatigue with the ongoing second scenario. The distinction of critical states did not allow a clear statement concerning satiation, which also increased from the first to the second scenario. While the sequence of dynamic density changing from high to low from the first to the second run still increased the cognitive impairments, a motivating and monotony-decreasing effect of the dynamic density changing from low to high was found. The monotony-decreasing effect of active exercises in rest breaks was confirmed, even though there was no favoring effect after repetitive conditions.

Boredom proneness and initial recovery and strain states were not confirmed to have a significant effect on monotony; marginally significant effects in univariate analysis do however indicate the relevance of further investigation. On the other hand, if individuals perceived flow during task execution in the first run, the indicator for the state of monotony was lower.

So far experimental results support the assumption that repetitiveness in task conditions is evoking a multidimensional individual response pattern as predicted by theory. It was discussed to further examine the contribution of traffic shifts as mitigation factor. The investigation of the monotony-relevant factors with a similar set-up in the operational environment was planned to come to general conclusions.

6. STUDY III: INVESTIGATION OF FACTORS IN AN OPERATIONAL ENVIRONMENT

6.1. OBJECTIVES AND HYPOTHESES

The main purpose of the study described in the following chapter was to investigate the issue of monotony and its contributing factors in an operational environment. Whilst the simulation-based experiment allowed defining essential variables under controlled conditions, these variables needed to be investigated in field settings to understand their relevance within the defined framework. The approach chosen to answer these questions consisted of the assessment of the controllers' psychophysiological state at defined work periods during their work shift. This included asking them how they felt and how they interpreted various aspects of the situation they were working in as well as assessing their physiological reactions. To gain information on behaviours, an observation at the working positions was combined with an extended interview at the end of each working session. Again, the objectives presented in Chapter 3 were addressed, even though from different perspectives.

Concerning objective 1, namely the investigation of task factors and the influence of individual and contextual factors on the development of a state of monotony, task factors were related to different requirements in the sectors. As shown in the laboratory studies, repetitive traffic situations cause suboptimal physiological activation and increase the subjective feeling of monotony and sleepiness. Impairments in concentration and attentiveness occurred in repetitive traffic situations even though less strain and mental workload were perceived, combined with longer conflict resolution times. Therefore, the objective was to find out how traffic characteristics relate to suboptimal states in field settings. Again, two different conditions were in the focus of interest. The concept of monotony was related to traffic flows that appear to be homogeneous on a long-term period. In addition, the effects of traffic load were investigated. It was tested if there was a difference in psychological functions during task execution depending on the number of aircraft. The goal was to define a critical number of aircraft.

The adaptation of the state to changing task demands was included to collect further evidence for the distinction of different critical states. How functional states develop after traffic peaks and how the performance impairments can be avoided was described.

Based on these preconditions, the following research hypotheses were investigated in this study related to the first research question in objective 1:

Hypothesis III.1$_C$[10]: There is a difference in a composed indicator for a state of monotony depending on traffic repetitiveness and traffic load in enroute air traffic controllers working at the executive position.

Hypothesis III.2$_C$: There is an influence of individual variables (initial state of strain and recovery, boredom proneness) on the composed indicator for the state of monotony.

Hypothesis III.3$_D$: There is a difference in various physiological, subjective and behavioural indicators and their development over time depending on the effect of traffic repetitiveness and traffic load.

Hypothesis III.4$_D$: There is an effect of traffic load on indicators of monotony and measures of cognitive functioning (e.g., attentiveness).

[10] C=confirmative hypotheses; D=descriptive hypotheses;

Objective 2 was focused on the description of monotony states in the field. One aspect was concerned with the description of different suboptimal functional controller states. In the simulation study it was shown that monotony not only changes as a function of repetitiveness, but also satiation increased during task execution and fatigue interacted with monotony after a certain time-on-task. Apart from that, the satiation indicators did not support the theoretically defined concepts. For this reason it is important to know how such a dynamic development of critical states can be described in the field, to make sure that countermeasures are introduced in time. Thus, a descriptive hypothesis was:

Hypothesis 5_D: There is a difference in indicators for critical states depending on time-on-shift.

In this way also the definition of warning signals that announce a decline in optimal controller states are of interest. Several signals were mentioned by controllers to announce a decline in vigilance, such as being surprised by aircraft or less scanning (Gordon, 2005). However, such signals as overt behaviour preceding inefficient states can be seen as warning signals and thus be useful in training. The definition of warning signals was addressed in an exploratory question to be answered in the interviews and determined through observation:

- Which characteristics precede and help to predict the development of critical states?

Finally, objective 3 was to define countermeasures. As it was found in the experimental study, physical break activity improves well-being and the perception of cognitive functioning. To see if a systematic position assignment depending on traffic predictions could be usefully employed in the field, a further research hypothesis investigated if changes in traffic density reported by the controllers have a different effect on monotony. In the lab study the switch from low to high traffic density led to increases in subjective motivation and a decrease in feeling of monotony. As an explorative issue the changes from low to high vs. high to low traffic density were considered to gain insight in the development of subsequent suboptimal states, as it was assumed to relate to alertness problems. In addition it was seen in the simulation study that shifts have also favourable effects on motivation and the reduction of monotony. The following additional hypothesis was addressed:

Hypothesis 6_D: There is a difference in motivation, concentration, and fatigue depending on the perception of changes in traffic density.

Another strategy that might have a big impact on safety-relevant issues but is hardly ever applied in the operational setting is the systematic variation in assignments to the planner and executive positions. One explorative question asked:

- Which effects does the systematic shifted exchange of planning controller (PC) and executive controller (EC) have on aspects related to performance and subjective well-being?

There were also a variety of additional strategies and their description was approached from two perspectives: the ones that controllers successfully used to maintain their optimal state respectively the ones they employed to improve an already suboptimal state. This concerned as well the factors that controllers applied to make their work interesting (e.g., creating variety in the task). Some of them might be rather obvious while others might occur rather unconscious (e.g., micro breaks) and thus can be discussed when reconstructing the work situation. Therefore the following explorative question was added:

- Which are the strategies used by controllers to maintain an optimal state or to improve a suboptimal state?

An additional aspect related to objective 1 is the investigation of effects of the night shift and was addressed in an additional design. Night shift also served to investigate the time-on-shift effect. The reason for the interest in night shifts is also its relation to repetitiveness. Controllers served as test subjects in the so-called WEST sector (presented in Figure 48), one sector combination predominantly activated in the night shifts. Traffic in this sector has a reputation of being repetitive and dense because of a lot of eastbound overflights. The sector was especially known to be repetitive later in the night shift, while rather non repetitive earlier in the shift. Therefore, an additional design was included for the separate analysis of work shift effects.

6.2. METHOD

6.2.1. Experimental Design

This field study was conducted in an east European control center, where the airspace could be structured in up to seven enroute sectors. In further references local labels are used to designate the sectors. Commonly, at the time of the study, sectors were split into low (L) and top (T) and EAST and WEST sectors. In periods of high traffic density, an additional division level was introduced to create middle (M) sectors. The hierarchy of sector collapsing procedures combined the lower and higher sectors to the EAST respectively WEST sector and one overall sector was frequently open during a certain period in the second part of the night shift. The study was based on a 2 (Repetitiveness) x 2 (Traffic Load) within-subjects design, resulting in four different conditions. The sectors and periods of interest were determined after asking four centre supervisors concerning the usual traffic density, complexity, and repetitiveness in commonly distinguished 90-minute-periods for each of the sectors during an average working day. Average scores were calculated on these periods and extreme groups (>+/-1 SD in each item) were considered for further validation with COSAAC (EUROCONTROL, 2005). This tool uses traffic data provided by the Central Flow Management Unit (CFMU) and allowed to cross-check the provided ratings of the supervisors. Sectors were selected which represented the expected combination of the two main factors. The low sectors were considered as non repetitive sectors, while the high sectors were considered to be repetitive.

The section within the work period was introduced as a further factor for selected dependent variables repeatedly measured during a work period. The 90-minute work periods were divided into three sections of 30 minutes each. The dependent variables comprised physiological and subjective indicators and additional influencing factors were considered that control or confound the results. Each of the ten air traffic controllers participated under four experimental conditions, presented in Table 47.

Table 47: Experimental within-subjects design (Study III)

		REPETITIVENESS	
	N=10	*non repetitive*	*repetitive*
TRAFFIC LOAD	*high*	Section 1 Section 2 Section 3 (n=10)	Section 1 Section 2 Section 3 (n=10)
	low	Section 1 Section 2 Section 3 (n=10)	Section 1 Section 2 Section 3 (n=10)

Table 48 contains an extended presentation of the experimental plan. Position assignments were counterbalanced for day 1 and day 2, controllers who worked according to schedule A on the first day worked on schedule B on the second day and *vice versa*. In schedule A the controller worked in sector WEST HIGH in the working period (WP) 1, and in EAST LOW in WP 2. In Schedule B the controller worked in WEST LOW in work period 1 and in WEST HIGH in work period 2.

Table 48: Experimental plan (Study III) for the assigned sectors on each study day exemplified for two participants

Study day	Day 1		Day 2		...
Participant	ATCO 1	ATCO 2	ATCO 1	ATCO 2	...
ATCO Schedule	A	B	B	A	...
Work Period 1	WEST HIGH	WEST LOW	WEST LOW	WEST HIGH	...
Work Period 2	EAST LOW	WEST HIGH	WEST HIGH	EAST LOW	...

Deviations from the initially planned design were necessary on the first study day due to unexpected sector collapsing procedures at the EAST sectors. The controller working in the EAST sector in the second work period was measured in the EAST LOW instead of the EAST HIGH. Thus, this resulted in an inconstant sector assignment between experimental conditions. One of the controllers was working at the WEST sector in both periods during one day, while the second one switched between the EAST and the WEST. However, on the level of the sector characteristics it was possible to compare both the EAST and WEST lower sector in terms of the usual requirements.

To control the impact of the first work period on the second one, a systematic rest break of 45 minutes was introduced before controllers sat down at the assigned position. Like this, an independent analysis of the conditions was possible. A further advantage of assessing two work periods at one day was to assess the time effect to be integrated in the interpretation, similar to the experiments in the laboratory setting.

The additional night shift design (Table 49) included the time-on-shift as well as the time on position as within-subjects factors. Each participant was measured during one night shift.

Table 49: Additional experimental within-subjects design night shift

<table>
<tr><th>N=10</th><th>TIME-ON-SHIFT</th><th></th></tr>
<tr><td rowspan="2"></td><td>Work period 1</td><td>Section 1
Section 2
Section 3
(n=10)</td></tr>
<tr><td>Work period 2</td><td>Section 1
Section 2
Section 3
(n=10)</td></tr>
</table>

6.2.1.1. Independent variables (IV)

<u>Traffic Repetitiveness (repetitive vs. non repetitive)</u>

Predominantly repetitive vs. non repetitive traffic periods in comparable sectors were selected. As already mentioned, it turned out that the low sectors demonstrated non repetitive characteristics whilst high sectors were considered to be repetitive at the selected periods.

Repetitiveness refers to the requirement of homogeneous and uniform working methods to deal with the traffic in a certain sector. It does not relate to the traffic itself, but to the activities of the controller and the variety of potential solutions. Thus, it is distinguished from sector complexity, an indicator clearly referring to static sector characteristics such as the number of crossing points, routes, climbing/descending traffic. The aspect of repetitiveness deviates from this definition as it is focused on the individual activity required to deal with the traffic (e.g., “hello-good-bye traffic”).

For example, a period of high traffic complexity may require complex or homogeneous/ repetitive working methods. Figure 48 shows characteristic traffic flows in both sectors. While the high sectors are characterized by homogeneous flows requiring simple actions, the low sectors are characterized by traffic requiring a variety of different actions.

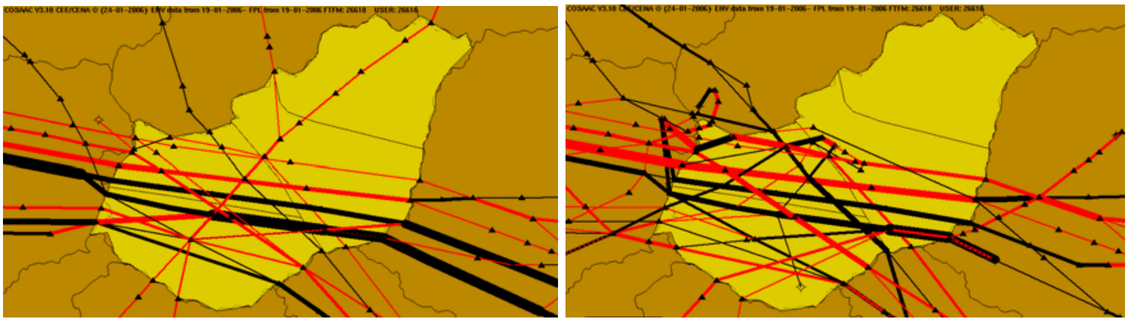


Figure 48: Traffic flow in high (left) and low (right) WEST sector in comparable traffic periods.

It might be argued that the sectors mainly distinguish in terms of sector complexity; this is however not focussed on in the study. It is rather homogeneity in traffic flows that was considered. Moreover, static sector complexity indicators such as the crossing points do not necessarily distinguish in the low and high sectors; it is rather the actual traffic that requires different actions. Also, there was more climbing and descending traffic in the lower sectors because of the vicinity of Vienna and Budapest airport, requiring a higher diversity of actions.

Traffic load (increased vs. decreased)

The second factor of interest was traffic load manipulated between high and low traffic load. The traffic load is an indicator similar to traffic density, with the difference that traffic density relates to sector volume. The sector volume of the chosen sectors was comparable. A different picture emerged, if the sectors were not collapsed.

The traffic load in the selected sectors showed a high variety. In all the measured sectors the average number of aircraft on frequency ranged from zero to 19 in a minute. Hence, this variation needed to be considered to relate indicators for states that might be used to specify bandwidth indicators for a reasonable workload management. Thus, the main experimental design depicted in Table 47 consisted of two variables with two manipulations on each factor, resulting in a 2 (repetitiveness) x 2 (traffic load) design with repeated measurements on each factor. For selected indicators (subjective ratings and physiological measures) additional measurements were collected in 3 vs. 9 sections during the work period.

The planned low density occurred systematically during the second period. The morning schedule which would have been characteristic for a low density period could not be chosen because of the needed preparation time. The potential after-effect from the first period was considered through a systematically planned rest break before the start of the second work period.

The planned variable traffic load could not be included as planned. This is because of local traffic characteristics, as not all supervisors decollapsed the sector EAST high and EAST low in the afternoon period. Therefore, an adaptation to the requirements was necessary and the traffic load manipulation of high and low is not constantly distributed between the first and the second work period. In case of collapsed sectors, these participants were excluded from statistical analysis in the relevant factors.

Time on position (3/9 sections)

Each controller was working at the sector for 90 minutes. Therefore, in the post monitoring questionnaire the indicators were asked for each 30-minute-period in the sector, that is section 1 (0-30 min) vs. section 2 (30-60 min) vs. section 3 (60-90min). In addition, heart rate indicators were analyzed for 10-minute-intervals.

6.2.1.2. Dependent variables (DV)

The controllers' state was assessed through a combination of physiological and subjective indicators.

Composed indicator for the state of monotony

The composed indicators of monotony was used with the standardized average inverted heart rate, average ratings of subjective feeling of monotony and sleepiness for each work period. [11]

Physiological indicators

Two types of indicators were measured, the electrocardiogram (ECG) and the blood pressure (BP). Automatic blood pressure monitoring (ABPM) was used for each participant on one day, which allowed a continuous comparison of the reactions of the cardiovascular system for selected work periods. BP was measured in 30-minute-intervals and was scheduled for 3 measurement periods after 15 minutes in each of the sections. The following indicators were assessed:

Table 50: Summary of physiological variables (Study III)

Level	Dependent Variable
ECG	Average heart rate in 10-minute-sections during run
	Average heart rate variability in 10-minute-intervals during run
	Average baseline of heart rate in 3-minute-intervals in the morning and after each work period
	Average baseline of heart rate variability in 3-minute-intervals in the morning and after each work period
Blood Pressure	Average baseline of systolic blood pressure in the morning and after each work period
	Average baseline of diastolic blood pressure in the morning and after each work period
	Average systolic blood pressure in each of the 3 sections of the work periods (restricted sample)
	Average diastolic blood pressure in each of the 3 sections of the work periods (restricted sample)

Subjective indicators

A variety of questionnaires was applied at several measurement points to assess the subjective reactions. In addition, how controllers perceive and interpret the situation was of a major relevance and expected to contribute to the understanding of the effect of complacency in relation to monotony. Complacency is a term that has been introduced in automation research (Parasuraman, Molloy, & Singh, 1993) and expresses a feeling of contentment and self-satisfaction. Because of a potential unawareness of danger, as one has the feeling to have everything under control, it may relate to repetitive traffic situations of low difficulty. The indicators measured are summarized in Table 51.

[11] The composed indicator as presented in Appendix A, Formula 1, and discussed in Chapter 4.4.3.1. was applied, as it is not possible to use several indicators in the mixed models procedure to repeat the principle of the MANOVA approach.

Table 51: Summary of psychological variables (Study III)

Level	Dependent Variable
Cognitive, emotional and motivational indicators (TSI)	flow, sleepiness, concentration, motivation, attention, boredom, feeling of monotony, confidence for 3 sections during work collected after work at position
Ratings of Traffic Characteristics	Traffic density, traffic complexity, traffic routine, traffic repetitiveness, traffic difficulty
Workload (NASA-TLX)	Mental, temporal demand, frustration, performance, effort for 3 sections during work collected after work at position
Mood (UWIST)	Average value for critical states (SOF) after work at position;
Critical States (SOF)	Average Values for mood (UWIST) after work at position;

Behavioural and performance indicators

This type of indicators combines the ratings of performance aspects after each work period and the SET-W rating on safety, efficiency, taskload, and workload in personally relevant situations.

Generally, a multi-level multi-method approach is even more meaningful if the subjective and physiological data is completed with behavioural information on performance. However, as it was assumed that in the operational environment controllers are usually providing their best performance, in the current study it was preferred to ask controllers if they personally noticed any deviations in their performance. One assumption is that behaviours mentioned by controllers to indicate low vigilance (e.g., missing calls) are related to a suboptimal individual state, where their ability to remain attentive and thus react appropriately is reduced.

6.2.1.3. Moderator and control variables

Various control variables were collected to describe the sample, namely biographical information on age, education, experience, handedness, body weight, body height, position assignments during the working day, break activities, initial state, and additional functions in the job. Traffic characteristics were collected after each work period to control the manipulation of the variables.

The moderator variables of interest are contained in Table 52. Additional information was collected which helped to support the interpretation of the data, such as behaviours during the work period at the sector. As it was not possible to collect and systematically analyze certain indicators during all work periods, the information was used to complete the picture.

Table 52: Summary of moderator variables (Study III)

Scale	Variable
RESTQ	Average scores in recovery Average scores in stress Average scores in subscales (General Stress, Emotional Stress, Social Strain, Conflicts, Overfatigue, Lack of Energy, Somatic Complaints, Success, Social Recovery, Somatic Recovery, General Recovery, Recovery Sleep)
ACS	Average Scores in Decision Related Action Orientation Scale (AOD) Average Scores in Action Orientation after Failure Scale (AOF) Average Scores in Action Orientation during Successful Performance (AOP)
MES	Average Scores in morningness-eveningness preference Morningness-eveningness types derived from MES
BPS	Average score in boredom proneness
IPIP	Personality Inventory
PAZ-K	Work satisfaction

Scale	Variable
Predictability	Average ratings of predictability before the beginning of the work period
SET-W	Ratings of Safety, Efficiency, Taskload – Workload at the beginning of the day
Traffic information	Average traffic load per hour Average aircraft under control on a minute to minute basis aggregated to 10 and 30 minutes.
Work period information	Shift from Low – High vs. High – Low
	PC exchange

6.2.2. Procedure

One European control centre agreed to host the study. For preparations, a day was arranged to discuss the planned sectors and procedures with the operational management, to observe on different sectors, and to provide a handbook with pre-study questionnaires and planned procedures for the volunteering controllers. The control center agreed to adapt the working schedules of the controllers participating in the study. General information was provided to the centre and the supervisors that the study took place. Similar to the simulator studies, controllers were instructed that the focus is to investigate the interesting aspects of their work creating variety.

The data collection was undertaken between 2nd February and 13th February 2006. Each work team provided two controllers for the study period who participated each on two day shifts and in one night shift. The shift supervisor received information on the schedules. To avoid any impact on safety, it was tried to avoid additional workload from the participation in the study. Since the interviews took place in the scheduled rest breaks, in some cases break times were reduced. This was not seen as a problem by the participants, as traffic during the winter season is reduced. The study leader was present during the day shift; during the night shift data collection was undertaken independently by the participants after prior instruction had occurred.

The experiment was carried out at the executive positions of the WEST LOW, WEST HIGH and EAST LOW sectors. A log book was prepared for each participant on each testing day that contained all the questionnaires. In the morning the scheduled positions were checked with the duty supervisors who were also informed about additional procedures. At the beginning of the first day in each team a short presentation was given to present the study goals and answer open questions. An extended briefing was undertaken with participants in a room provided by the center. Data was collected in different conditions from each person according to the experimental plan.

The multi-method approach considered the assessment of the controller's state in four different working situations during the day shift. In addition, recordings are available for each controller during two work periods in the first half of the night shift. Each work period lasted for 90 minutes. In Table 53 the procedure of a study day is described. A participant was randomly assigned to Schedule A or Schedule B during one day shift, and the schedules were exchanged for the second day shift. At the beginning of a shift the controllers were scheduled for preparation of physiological measurements and to answer pre-shift-questionnaires. After completion of the work periods (WP) at the sectors, interviews were conducted.

Table 53: Standard procedure of one study day for each work schedule (Study III)

Day shift (UTC)				Night shift	
Time (UTC)	Schedule A	Schedule B		Schedule C	Time (UTC)
6:00 – 6:15	Controller Briefing			Controller Briefing	18:00 – 18:15
6:15 – 6:20	Short Presentation (5 min) of the study				
6:20 – 6:45	Preparation of the study participants: prepare heart rate recorder; fill-in pre-shift questionnaires			Preparation of the study participant & fill-in pre-shift questionnaires	18:15 – 18:40

Day shift (UTC)				**Night shift**	
Time (UTC)	**Schedule A**	**Schedule B**		**Schedule C**	**Time (UTC)**
6:45 – 10:30	Assigned tasks or rest breaks by the supervisors				…
10:30 – 12.00	Work Period 1 WEST HIGH *(repetitive/ high traffic load condition)*	Work Period 1 WEST LOW *(non repetitive / high traffic load condition)*		Period 1 Work at EC in sector WEST	19:30 – 21:00
12:00 – 12:30	Fill in questionnaires and interviews (time for Lunch to be discussed with ATCO)			Fill in questionnaires	21:00 – 21:15
12:30 – 14:45	Assigned tasks or rest breaks by the supervisors				…
14:45 – 15:00	Rest break			Rest break	21: 45 – 22:30
15:00 – 16:30	Work Period 2 EAST LOW *(non repetitive / low traffic load condition)* *Planned:* *EAST HIGH* *(repetitive / low traffic load)*	Work Period 2 WEST HIGH *(repetitive/low traffic load condition)* *Planned:* *WEST HIGH* *non repetitive / low traffic load*		Period 2 Work at EC in sector WEST	
16:30 – 17:00	Fill in questionnaires and interviews			Fill in questionnaires	22:30 – 22:45
17:00 – 18:00	Assigned tasks or rest breaks by the supervisors				…

6.2.3. Participants

A sample size of 10 was determined to be sufficient to detect effects in a complete within-subjects manipulation: power was set to be at least 80 percent, with an alpha level of 5 percent and beta error set to 20 percent. As seen in the simulation study, high effect sizes could be expected. The required sample size was estimated with Erdfelder et al. (1996) and revealed that with n=10 a power of 0.88 could be expected for the main effects ($df_{num}=1$; $df_{denum}=9$; m=4). Participants were recruited through the assistance of the operational management. All subjects were selected between the age of 20 and 40 years. No female controllers could be included to equalize the field study with the simulator study, as the percentage of female controllers at the center was different. Supervisors were pre-informed concerning the required schedules.

The shift organization in the center consisted of one day shift from 7:00 to 19:00 followed by one night shift from 19:00 to 7:00 and two days off. The controllers are organized in 5 teams that alternately cover the shifts. Table 54 gives an example for the shift organization. For organizational reasons, two controllers from each team participated in two day shifts at two selected periods. Each night shift following the day shift one of the controllers participated. Randomization was considered in the assignment of the participants to the schedules, its possibilities were however limited. None of the selected participants withdrew from the study.

Table 54: Shift Organization Schedule

	Day 1	**Day 2**	**Day 3**	**Day 4**	**Day 5**
Day 7:00 – 19:00	TEAM A	TEAM B	TEAM C	TEAM D	TEAM E
Night 19:00 – 7:00	TEAM E	TEAM A	TEAM B	TEAM C	TEAM D

The volunteer's age ranged from 27 to 40 years with an average of 36 years. At the time of the study they had been licensed as air traffic controllers for a period of between two and 20 years. Half of the participants were engaged in work related projects and two of them have been active instructors, one for 5, one for 15 years. An overview of the descriptive statistics for the sample is presented in Table 55.

Table 55: Sample description of quantitative indicators (Study III)

	Mean	SD	min	max
Age	35,8	4,5	27	40
Years in Company	13,5	6,4	5	20
License in years	12,3	5,6	2	20
Weight	83,6	9,6	70	86
Height	175,6	4,9	170	182

Note. N=10

Frequencies of additional variables are presented in Table 56.

Table 56: Sample description of qualitative indicators (Study III)

Variable	Category	Frequency
Vision	normal glasses	9 1
Handedness	Right Both Left	8 1 1
Additional Tasks	Yes No	5 5
Instructor	Yes No	2 8

Note. N=10

6.2.4. Materials and Apparatus

6.2.4.1. Subjective indicators

The following table (Table 57) gives an overview of the questionnaires and scales which are further described subsequently. Most of the methods have already been described in Chapter 4 and Chapter 5. For this reason, only additional questionnaires are explained.

The predictability indicator was composed from the estimated traffic complexity and density and the number of estimated potential conflict situations in the subsequent work period. It was introduced to express how controllers react in case the traffic developed differently than expected, since it might influence how a controller adapts his state to required performance.

The post-monitoring scale was combined from several rating scales previously applied in study I and II (NASA-TLX, items used by Thackray et al., 1975) and thus deviates from the original scales to avoid inhomogeneous answer formats. The items were completed with ratings for traffic characteristics (difficulty, density, repetitiveness, routine, complexity) and answered separately for the three 30-minute-sections of each work period. Two items were added to ask for flow experience, based on Harris (2000). The last section of the questionnaire addressed the perceived performance impairments and specifically asked for how behaviors were affected. The 7-point-answer format of all integrated scales was identical to avoid confusion, the labels were adapted.

Table 57: Summary of materials and apparatus (Study III)

ADMINISTRATION	NAME (ABBR.)	AUTHOR
Before participation	Biographic questionnaire (BIO)	
	Personality Inventory (IPIP)	Goldberg 1999a

	Morningness Eveningness Questionnaire (MEQ)	Horne and Ostberg, 1976
	Action Control Style ACS 90	Kuhl, 1994b
	Boredom Proneness Scale BPS	Farmer and Sundberg, 1986
Beginning of study day	Initial State Questionnaire (ISQ)	Translation based on Janke, 1976
	Questionnaire for Strain and Recovery State (RESTQ)	Kallus, 1995
Before work period	Predictability rating	
After work period	Post-monitoring scale	Based on NASA TLX, items used by Thackray et al. (1975), Harris (2000) and additional items
	Thackray Scale Inventory (TSI)	translated version of BMS (Richter & Plath, 1984) by Rockstuhl, 2001;
	Mood Assessment UWIST	Matthews, Jones & Chamberlain, 1990
	SET/W-rating	Vormayr, Kallus & Hoffmann 2005
End of day	Debriefing guide	

6.2.4.2. Electrocardiogramm

Heart rate was recorded with a portable device (BHL 6000, Mednatic Munich). Additional devices required were a standard notebook from HP and the software package provided for the system. The recorder was programmed to record each heartbeat and adhesive electrodes were attached in the recommended standard position on the chest. After the recording, individual heart rates as well as their variation were determined. All continuously assessed measures were analyzed in 10 minute intervals and aggregated to 30 minutes or 90-minute-periods.

6.2.4.3. Blood pressure measurement

Blood pressure was assessed with two procedures. Every day, one controller participated in an automatic blood pressure measurement (ABDM), which was executed with the automatic blood pressure meter Cardio Tens CT1 (Meditech, HU) under local support by the medical team. ABPM was applied for eight participants during one working day. Two participants did no accept the device, as it was too interfering for them.The meter was fixed for a day and automatically inflated at pre-programmed intervals. In addition, a 3-minute-baseline recording was taken for physiological indicators at the beginning of the day as well as after the work at the positions. The blood pressure was analyzed by the local medical team with *Medibase* monitoring software and anonymously provided for the statistical analysis. Alternatively, blood pressure of the second participant was recorded with the BOSO MEDICUS PC, a BP meter manufactured by Bosch, and recorded in a data sheet.

6.2.4.4. Work satisfaction profile (PAZ)

This profile was developed by Jimenez (2000) to determine work satisfaction. The English version contains the subscales presented in Table 58.

Table 58: Subscales of the work satisfaction profile (PAZ)

Item	Abbrev.	Name	No. of Items
4-6	PHA	Satisfaction with how demanding your job is	3
7-10	PUK	Satisfaction with the contact to your colleagues	4
11-14	PBV	Satisfaction with the relationship to your nearest boss	4
15-17	POF	Satisfaction with organisation and management	3
18-20	PAB	Satisfaction with the working conditions	3
21-23	PES	Satisfaction with my freedom in decision-making	3
24-26	PBZ	Satisfaction with the payment I receive	3
27-31	PAU	Satisfaction with working- and vacation times	5
32-34	PAR	Satisfaction with the general working conditions	3

A final item was included to ask for the overall satisfaction with the work. Moreover, the relevance of various features and expected changes of the features are included in the assessment. The reliability and validity of this scale correspond to commonly accepted criteria. The internal consistence for the subscales range between r=.65 and r=.94, with a median of r=.90.

6.2.4.5. Observations

Observations did not only represent a preparation of reconstruction interviews, but helped to gain insight in the ongoing activities at the positions. Each controller was observed once each test day on selected positions. To follow the communication, headphones were provided. The scope was to get a picture of the variety of the sectors at the different work periods. Information was collected on special occurrences, aircraft under control in 10-minute-intervals, strategies to counteract, or signals that do announce changes in critical states such as increased yawning. The data helps to better understand physiological reactions. Even though the current observation schedule does not allow a comparison of all the situations, it was possible to have complementary information on the distribution of the traffic and how sectors were split.

6.2.4.6. SET-W (Safety-Efficiency-Taskload-Workload) - Rating

SET-W is a computerized rating scale (Kallus, Hoffmann, Ehgartner, Kuhn, Pichler, & Schuen-Medwed, 2003; Vormayr, Kallus & Hoffmann, 2005) to assess the safety, efficiency, taskload and workload of defined situations on a 5-point-scale ranging from -2 (low) to +2 (high). Overall, the balance of the indicators is evaluated. The rating was introduced at the beginning of the shift to assess the predictions for the day as well as for situations occurring during a work period that were perceived as personally relevant in terms of safety.

6.2.4.7. Combined reconstruction interviews and debriefing

An interview was conducted after the work periods and recorded on tapes. Initially planned to be executed after each working period, the local occurrences required changing this plan and it was tried to have at least one interview with each participant and to do additional interviews in case of critical situations. Part of the questions were based on the guide used in the lab studies and further interview questions were posed to determine subjective interpretations of issues already asked for in the post-monitoring scale, as related to performance, motivation, emotion, flow, cognitive processes, and the role of expectations, complacency, and over-confidence. Additional topics addressed the differences in work strategy, the influence of task-irrelevant thoughts, and the effects of traffic shifts and helped to answer the research questions. A major point was the definition of warning signs for a critical state and strategies to improve a critical state.

6.2.5. Data Processing

The data processing was already described in 6.4. This chapter only contains additional procedures or deviations. Statistical procedures were performed with SPSS 14.

The following situations from the actual procedure were considered in data analysis. Concerning the effects of repetitiveness and traffic density in the mixed model the data of three participants was excluded during one work period each due to sectors collapsed by supervisors and thus having an impact on the traffic demands in the sector. This was necessary to guarantee the homogeneity and comparable conditions. For the comparison of the traffic load independently from repetitiveness in sectors all data was included.

Missing data in physiological indicators were due to system failure, resulting in a loss of data for ECG in 5 percent. Two persons did not accept the ABPM, one person did not fill in two pre-study questionnaires and a subjective scale. The advantage of the selected mixed models procedure was also that missing data did not result in loosing the case that would be fully excluded if one datum was missing.

The interviews were transcribed and the analysis of the interviews was undertaken based on the categories defined in the objectives and the research questions. As reported by Kallus et al. (1998) interviews with controllers work in English or German and need not be conducted in the native language of the controller or interviewer. For this reason, the demand of one controller to do the interview in German rather then English was accepted.

6.2.6. Statistical Analysis and Hypothesis Testing

The research hypotheses were already introduced in the beginning of this chapter and transformed to statistical hypotheses which could be submitted to statistical analysis. The main statistical hypotheses addressed the effects of the independent variables repetitiveness and traffic load on a composed indicator for the state of monotony. A further statistical hypothesis tested the influence of moderator variables included as an additional factor. These hypotheses were considered as confirmative as they were supposed to help deciding on the relevance of these factors in a model of monotony. The alpha level was set to $p=0.05$, the correction method of Bonferroni-Holm was applied to avoid alpha-inflation.

The additional statistical hypotheses addressed task factors and repeated measurement factors such as section during work period as independent variables and subjective or physiological indicators according to the listed variables in 6.2.1.2 as dependent variables. They were evaluated according to the descriptive analysis method of Abt (1983). The recommended level was maintained as the consequences of making a false decision in favour of H1 is judged as more critical, since the consideration of additional task factors in the operational environment requires to put forward initiatives for their consideration.

The linear mixed model approach (Wallace & Green, 2002) was used for statistical analysis to evaluate differences in the assessed variables depending on traffic conditions. Traffic repetitiveness and traffic load indicators were considered as independent variables and an index variable marked the repeated measurements within each participant. A model was determined for each dependent variable. The mixed models approach requires the definition of the appropriate covariance structure through assessing the best fitting model for the data structure, where the null hypothesis tests that a smaller model with less parameters provides as good a fit for the data than the larger model. The models were determined separately for the confirmative hypothesis; the dominating structure was maintained in additional variables.

6.3. RESULTS

Detailed information concerning descriptive indicators such as mean values (M) and standard deviations (SD) and results of the statistical analysis not reported in detail are included in Appendix C.

6.3.1. Confirmative Hypotheses 1: Effects of Repetitiveness and Traffic Load on Monotony

To test the effect of repetitiveness and traffic load on the standardized indicator for the state of monotony, a statistical analysis compared a total of the 90-minute-work periods in the different sectors. The analysis using a linear mixed models approach included traffic load and repetitiveness as factors with an index marking the repeated measurements for each participant.

As in previous studies, the indicator for the state of monotony was composed[12] with the z-scores of ratings of feeling of monotony, sleepiness and the inverted heart rate in each working period to standardize the units (descriptive statistics in Table C-1). All the variables were normally distributed after conducting the Kolmogorov-Smirnov-Test (Monotony Indicator: Z=.83, p=.498; Feeling of Monotony: Z=.82, p=.507; Sleepiness: Z=1.01, p=.261; HR inv. Z=.61, p=.849). For the analysis of the composed indicator for the state of monotony, a comparison of the models for the different covariance structures was conducted submitting the number of model parameters and magnitude of the information criteria to a Chi²-test. Compared with the model for unstructured variances, the model for compound symmetry was superior for the indicator for the state of monotony (Model parameter 9, 8; p<.10). While for sleepiness the compound symmetry model had a better fit for the covariance structure (model parameter 11, 8; p>.10), the unstructured model was more appropriate in ratings of subjective feeling of monotony and inverted HR (model parameter 15, 8; p<.10; model parameter 26, 8; p<.10).

The analysis revealed a significant effect of repetitiveness ($F_{1,25}$=8.66, p=.007) and a marginally significant interaction between traffic load and repetitiveness ($F_{1,25}$=3.45, p=.075), the effect of traffic load was not significant ($F_{1,25}$=1.27, p=.271). The indicator for the state of monotony is higher in repetitive conditions. As summarized in Table 61, the null-hypothesis addressing the effect of repetitiveness is rejected and the alternative hypothesis assumed. The null-hypothesis concerning the main effect of traffic density on the indicator of monotony was retained. Further univariate analysis revealed that the indicator of monotony was significantly higher in the repetitive low traffic load condition ($F_{1,25}$=4.63, p=.045).

[12] It was not possible to apply MANOVA procedure, therefore the composed indicator was chosen to apply the Mixed Models procedure.

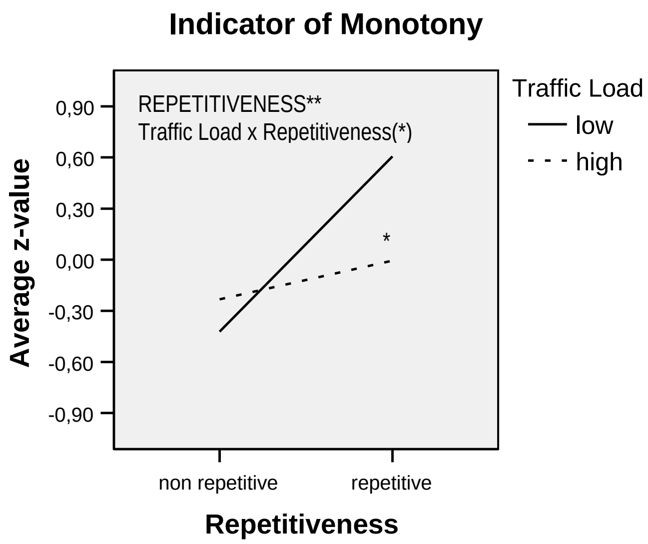


Figure 49: Average values for the composed indicator of monotony as a function of repetitiveness and traffic load.

Further comparison of the groups depicted in Figure 49 resulted in a significant difference between increased and reduced traffic load only under the repetitive condition. A separate analysis for the single indicators showed a clear difference in the repetitive conditions in the individual ratings of sleepiness and the subjective feeling of monotony for that period (Table 59).

Table 59: Results of Linear Mixed Model Analysis for single indicators of monotony (Study III)

Source	Results ($F_{df\ nominator,\ df\ denominator}$, p-value)		
	HR inv. (UN)	Sleepiness (CS)	Feeling of Monotony (UN)
Repetitiveness	$F_{1,9}$=.33; p=.580	**$F_{1,25}$=6.865, p=.015***	**$F_{1,0.33}$=67.91, p=.002****
Traffic Load	**$F_{1,118}$=129,7; p=.000*****	$F_{1,25}$=2.171, p=.153	$F_{1,132}$=.30, p=.602
Repetitiveness x Traffic Load	**$F_{1,9}$=14.2; p=.004****	$F_{1,25}$=1.186, p=.286	**$F_{1,14}$=3.66, p=.088(*)**

*Note. 3 missing work periods. ***p<.001. **p<.01. *p<.05. (*)p<0.1. Covariance Structure of Mixed Models: UN=unstructured variances, CS=compound symmetry.*

No effect of repetitiveness was found in heart rate, there was however a significant effect and interaction with traffic load. Heart rate was rather affected by the traffic load.

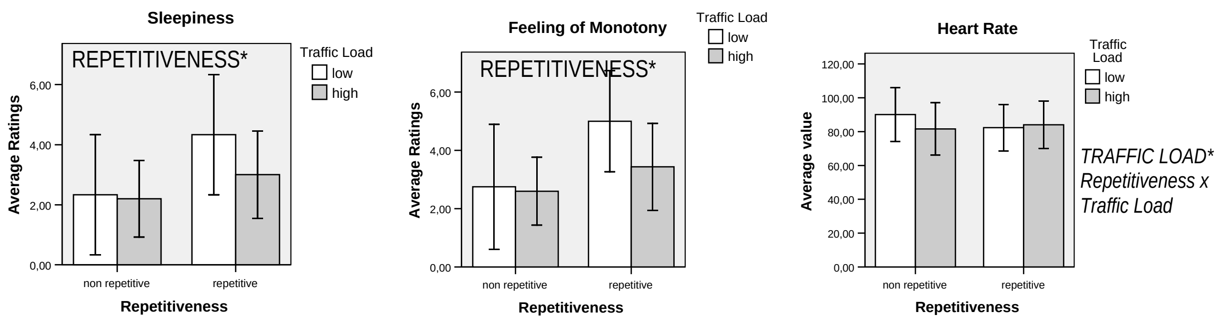


Figure 50: Average values for non-corrected ratings of sleepiness, subjective feeling of monotony and heart rate as a function of repetitiveness and traffic load.

6.3.2. Confirmative Hypotheses 2: Influence of Individual Factors on Monotony

Different factors have been analyzed with respect to assess their influence on the states of monotony. They were expected to have an impact on how individual states develop during the day. Thus, analysis of co-variance and blocked designs - if ANCOVA was not possible - were included for these indicators.

Table 60: Results of Linear Mixed Model Analysis for moderator variables (Study III)

Source	Results ($F_{df\ nominator,\ df\ denominator}$, p-value)		
	Boredom Proneness (CS)	Strain (CS)	Recovery (CS)
Blocking Factor	$F_{1,8}$=.19; p=.680	$F_{1,8}$=1.15, p=.314	$F_{1,7}$=14.32; p=.006**
Repetitiveness	$F_{1,22}$=5.70; p=.026*	$F_{1,25}$=8.82, p=.007**	$F_{1,25}$=8.45; p=.008**
Traffic Load	$F_{1,21}$=2.04; p=.168	$F_{1,25}$=1.32, p=.262	$F_{1,25}$=1.66; p=.210
Repetitiveness x Traffic Load	$F_{1,22}$=2.15; p=.157	$F_{1,25}$=3.54, p=.072(*)	$F_{1,25}$=3.30; p=.081

*Note. 3 missing work periods. ***p<.001. **p<.01. *p<.05. (*)p<0.1.*

The initial states of strain and recovery were assessed with the Recovery-Stress-Questionnaire (RESTQ, Kallus, 1995). The inclusion of recovery state as a blocking variable revealed a significant effect of the initial state of recovery (Table 61). The mixed linear model was applied with compound symmetry covariance structure (model parameters 9,8; p>.10). Participants who felt less recovered at the beginning of the working day also showed a higher value in the monotony indicator (Table 60). There was however no effect of boredom proneness or initial state of strain (descriptive statistics Table C-2). As summarized in Table 61, the alternative hypothesis can be assumed for the state of recovery, while the null hypothesis for boredom proneness and initial brain state are maintained.

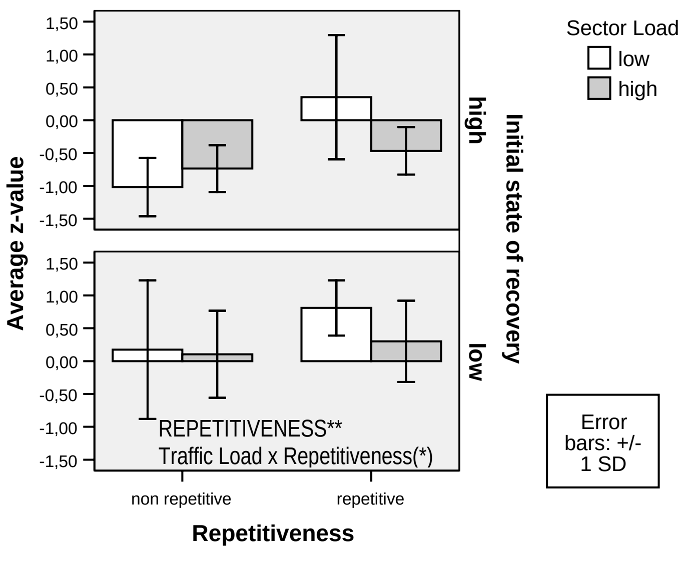


Figure 51: Average standardized values and SD of the indicator for monotony as a function of repetitiveness and traffic load depending on high and low initial recovery.

6.3.3. Testing of Confirmative Hypotheses

Again, to evaluate the null hypothesis, the correction of the alpha level was undertaken according to the procedure proposed by Bonferroni-Holm. For this reason, all p-values for the main effects were sorted starting from the smallest one.

Table 61 contains the confirmative hypothesis and the result for the main effects. The null hypotheses concerning the effect of boredom proneness and initial strain were retained. The null hypothesis for the effect of monotony and influence of state of recovery were rejected. To further come to a decision for the alternative hypothesis, the significant interaction effects were assessed. Because of the non-significant difference in the non repetitive condition, the alternative hypothesis was assumed.

Table 61: Correction of alpha level for the confirmative hypotheses and related decisions (Study III)

Confirmative Hypothesis	Description	p-value	Rank	Adjusted alpha level	Decision for H0	Decision for H1
H1.1	Main effect of repetitiveness on monotony	.007	2	0.025	rejected	assumed
H1.2	Main effect of traffic load for monotony	>.05	-	-	retained	rejected
H2.1	Influence of boredom proneness	>.05	-	-	retained	rejected
H2.2	Influence of state of recovery	.006	1	0.05:2=0.025	rejected	assumed
H2.3	Influence of state of strain	>.05	-	-	retained	rejected

6.3.4. Description of Additional Results

6.3.4.1. Effects of task characteristics on subjective, behavioral and physiological indicators (Descriptive Hypotheses 3)

The subjective and physiological indicators were compared with the mixed models procedure for the total of the 90-minute-work periods in the different sectors. Repetitiveness and task load were included as factors and interaction effects were determined for each indicator of interest. As CS turned out to better represent the covariance structures and required less parameters, this model was maintained for further descriptive analyses. In case of significant deviations requiring a different variance structure, this is noted. Only deviations from the assumption of normally distributed variables are reported.

The results of the statistical analysis for subjective item list based on Thackray and additional indicators for flow and confidence are summarized in Table 62, descriptive statistics are contained in Table C-3.

Table 62: Results of Linear Mixed Model Analysis for individual ratings (Study III)

Source	Results ($F_{df\ nominator,\ df\ denominator}$, p-value)						
	Attentiveness	Strain	Concentration	Boredom	Motivation	Flow	Confidence
Repetitiveness	**$F_{1,27}$=6.65 p=.016***	**$F_{1,25}$=12.89 p=.001*** **	**$F_{1,27}$=16.13 p=.000*** **	**$F_{1,25}$=7.48 p=.011***	***$F_{1,25}$=3.03 p=.094(*)***	**$F_{1,24}$=1.30 p=.266**	$F_{1,24}$=1.30 p=.265
Traffic Load	$F_{1,27}$=1.09 p=.306	$F_{1,25}$=2.67 p=.115	$F_{1,26}$=1.74 p=.198	$F_{1,25}$=3.78 p=.063	**$F_{1,25}$=8.03 p=.009** **	$F_{1,24}$=1.01 p=.324	**$F_{1,24}$=3.27 p=.083(*)**
Repetitiveness x Traffic Load	$F_{1,27}$=.12 p=.737	$F_{1,25}$=.00 p=.977	$F_{1,27}$=1.05 p=.315	$F_{1,25}$=1.41 p=.247	$F_{1,25}$=.14 p=.708	$F_{1,24}$=1.61 p=.216	**$F_{1,24}$=10.39 p=.004** **

*Note. 3 missing work periods. ***p<.001. **p<.01. *p<.05. (*)p<0.1.*

The main effect of repetitiveness was significant in the ratings of attentiveness, strain, concentration and boredom, while the effect of traffic load was significant in the ratings of motivation and marginally significant in the boredom rating. The controllers felt less attentive and concentrated, but more bored and strained in the repetitive condition. Motivation was rated lower under reduced traffic load. Figure 52 depicts differences in a selection of ratings.

Figure 52: Average ratings and SD for attentiveness, boredom and motivation as a function of repetitiveness and traffic load.

Interestingly, further univariate analysis of the significant interaction between repetitiveness and traffic load in the perceived confidence in the work revealed (Figure 53) that in the repetitive condition monotony was rated significantly higher under low traffic load ($F_{1,24}$=13.20, p=.001). There was however no significant effect in the items introduced to assess flow.

The results of the statistical analysis for the individual and composed indicators of the NASA-TLX are contained in Table 63 (descriptive statistics in Table C-4). As depicted in Figure 53, the workload was significantly lower in repetitive conditions and under reduced traffic load. A similar development was reflected in the ratings of mental demand, temporal demand and effort. Controllers felt to perform better in the non repetitive sector under reduced traffic load.

Table 63: Results of Linear Mixed Model Analysis for subjective ratings (Study III)

Source	Results ($F_{\text{df nominator, df denominator}}$, p-value)					
	Workload	Mental	Temporal	Effort	Performance	Frustration
Repetitiveness	**$F_{1,26}$=17.05 p=.000*****	***$F_{1,8}$=4.62 p=.062(*)***	**$F_{1,7}$=23.72 p=.001****	**$F_{1,8}$=6.52 p=.033***	**$F_{1,9}$=6.92 p=.027***	$F_{1,9}$=.03 p=.856
Traffic Load	**$F_{1,26}$=6.32 p=.019***	**$F_{1,8}$=7.02 p=.030***	**$F_{1,7}$=5.83 p=.049***	**$F_{1,5}$=18.44 p=.007****	$F_{1,9}$=.80 p=.394	$F_{1,8}$=2.35 p=.160
Repetitiveness x Traffic Load	$F_{1,26}$=.13 p=.722	$F_{1,8}$=.08 p=.779	$F_{1,8}$=3.33 p=.113	$F_{1,7}$=.17 p=.696	***$F_{1,9}$=5.00 p=.052(*)***	$F_{1,5}$=3.87 p=.109

*Note. 3 missing work periods. ***p<.001. **p<.01. *p<.05. (*)p<0.1.*

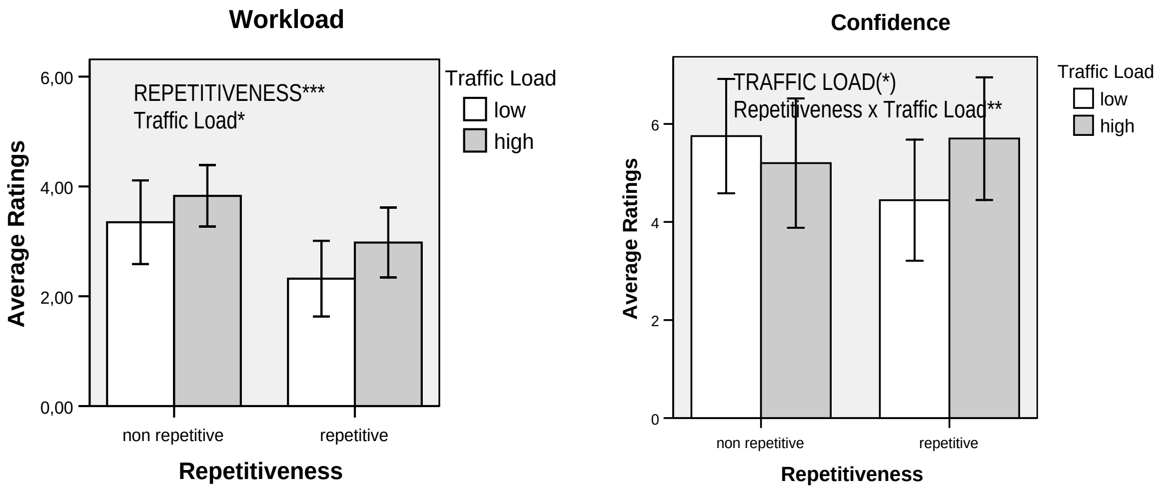


Figure 53: Average ratings and SD of workload and confidence as a function of repetitiveness and traffic load.

For the measurement of critical controller states after work at the sectors the *scale of feelings* (SOF) was administered to assess monotony, fatigue, satiation, and stress. The results of the mixed model analysis are presented in Table 64 (descriptive statistics in Table C-5).

Table 64: Results of Linear Mixed Model Analysis for scale of feeling (SOF) subscales (Study III)

Source	Results ($F_{df\ nominator,\ df\ denominator}$, p-value)			
	Satiation	Fatigue	Monotony	Stress
Repetitiveness	$F_{1,23}$=.26;p=.615	$F_{1,23}$=.24;p=.632	$F_{1,23}$=.00;p=.978	$F_{1,23}$=1.03;p=.320
Traffic Load	$F_{1,23}$=2.25;p=.147	$F_{1,23}$=.75;p=.395	$F_{1,23}$=2.55;p=.124	$F_{1,23}$=.54;p=.470
Repetitiveness x Traffic Load	$F_{1,23}$=.64;p=.432	$F_{1,23}$=2.59;p=.121	$F_{1,23}$=2.97;p=.099(*)	$F_{1,23}$=.01;p=.936

*Note. 3 missing work periods. ***p<.001. **p<.01. *p<.05. (*)p<0.1.*

The interaction between repetitiveness and traffic load in the subscale for monotony was marginally significant, but no additional effects were found. Further univariate analysis to interpret the interaction revealed that under the repetitive condition the scores are significantly higher in low traffic load ($F_{1,23}$=5.52, p=.028) and is depicted in Figure 54.

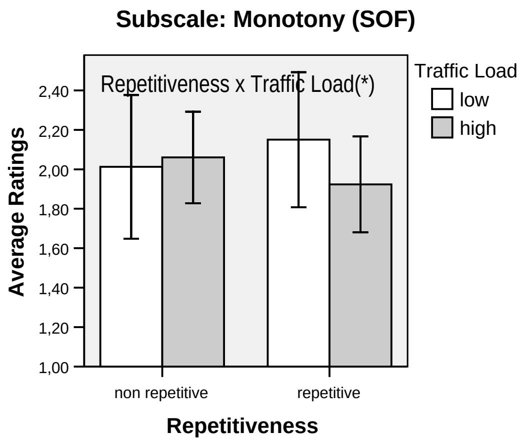


Figure 54: Average ratings and SD for the Scale of Feelings (SOF) subscale of monotony as a function of repetitiveness and traffic load.

Also, no significant effects were found in the UWIST mood assessment subscales (Matthews et al., 1990) for hedonic tone (Traffic Load: $F_{1,23}$=.89,p=.356; Repetitiveness: $F_{1,22}$=1.16, p=.294; Repetitiveness x Traffic Load: $F_{1,22}$=.11, p=.742), tense arousal (Traffic Load: $F_{1,23}$=.95,p.=339; Repetitiveness: $F_{1,22}$=.28, p=.605; Repetitiveness x Traffic Load: $F_{1,22}$=.69, p=.414) and energetical arousal (Traffic Load: $F_{1,24}$=.10,p=.752; Repetitiveness: $F_{1,23}$=1.49, p=.234; Repetitiveness x Traffic Load: $F_{1,23}$=.30, p=.589).

On the physiological level, no significant effects of traffic repetitiveness were found in HRV (Traffic Load: $F_{1,19}$=.94, p=.345; Repetitiveness: $F_{1,19}$=.45,p=.512; Repetitiveness x Traffic Load: $F_{1,19}$=1.46, p=.243). This is similar to the results of the analysis of heart rate. Concerning the development of the blood pressure in relation to the baseline, no significant effects of the task characteristics were found in systolic blood pressure (Traffic Load: $F_{1,24}$=.00, p=.998; Repetitiveness: $_{1,24}$=.07, p=.795; Traffic Load x Repetitiveness $F_{1,24}$=.03, p=.857).

There was a marginally significant effect of repetitiveness in the diastolic blood pressure ($F_{1,24}$=3.02, p=.095). The baseline-corrected diastolic blood pressure increased higher in the low traffic conditions compared to the high traffic conditions (Figure 55). No further effects were found (Traffic Load: $F_{1,24}$=2.70; p=.113; Repetitiveness x Traffic Load: $F_{1,24}$=.25; p=.624). During the work periods, systolic and diastolic blood pressure was only measured with the ABPM for a restricted sample (n=8) during three sections for two work periods at one study day. The test participants were working in the repetitive sectors during work period one and the non-repetitive sector during work period 2. In systolic and diastolic blood pressure, no significant effects of time on position or between the two types of sectors were found (SBP: repetitiveness: $F_{1,7}$=1.46, p=.265; time on position: $F_{2,7}$=.64, p=.558; DSP: repetitiveness: $F_{1,7}$=.04, p=.855; time on position: $F_{2,7}$=1.48, p=.316). A significant interaction was however found between time on position and the repetitiveness in sectors in systolic blood pressure ($F_{2,7}$=5.79, p=.033; presented in Figure 55), but not in diastolic blood pressure ($F_{2,7}$=2.92, p=.121). Further univariate analysis revealed that this effect is mainly reflected in the second section.

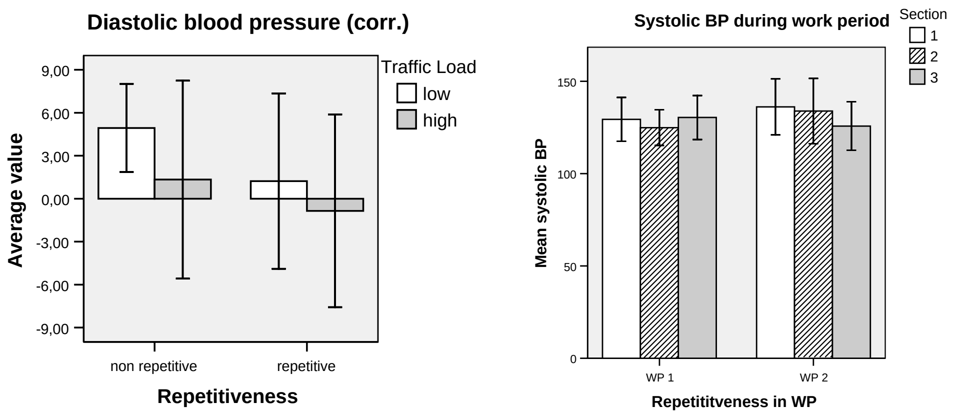


Figure 55: Average values and SD for baseline-corrected blood pressure as a function of repetitiveness and traffic load (left graph) and for systolic blood pressure during both work periods (WP) (right graph; n=8).

A further aspect concerned the controllers' performance, which was mainly addressed with self-assessment of behaviours on a scale ranging from 0 (never) to 6 (always). As depicted in Figure 56, ATCOs reported significantly less scanning in repetitive conditions (Repetitiveness: $F_{1,24}$=6.84, p=.015; Traffic Load: $F_{1,24}$=.15, p=.698; Repetitiveness x Traffic Load: $F_{1,24}$=.01, p=.920) and felt more easily distracted (Repetitiveness: $F_{1,22}$=.4.71, p=.041; Traffic Load: $F_{1,22}$=.87, p=.361; Repetitiveness x Traffic Load: $F_{1,22}$=1.27, p=.272). They also reported to feel less focused in repetitive conditions (Repetitiveness: $F_{1,24}$=4.75, p=.065; Traffic Load: $F_{1,24}$=.09, p=.764; Repetitiveness x Traffic Load: $F_{1,24}$=.01, p=.904), but thought to react slower under low traffic load (Repetitiveness: $F_{1,24}$=2.80, p=.107; Traffic Load: $F_{1,24}$=3.77, p=.064; Repetitiveness x Traffic Load:

$F_{1,24}$=.68, p=.419); these effects were however marginally significant. No effects were found in any of the other indicators, thus the details of the statistical analysis and descriptive statistics are contained in Table C-6 and C-7.

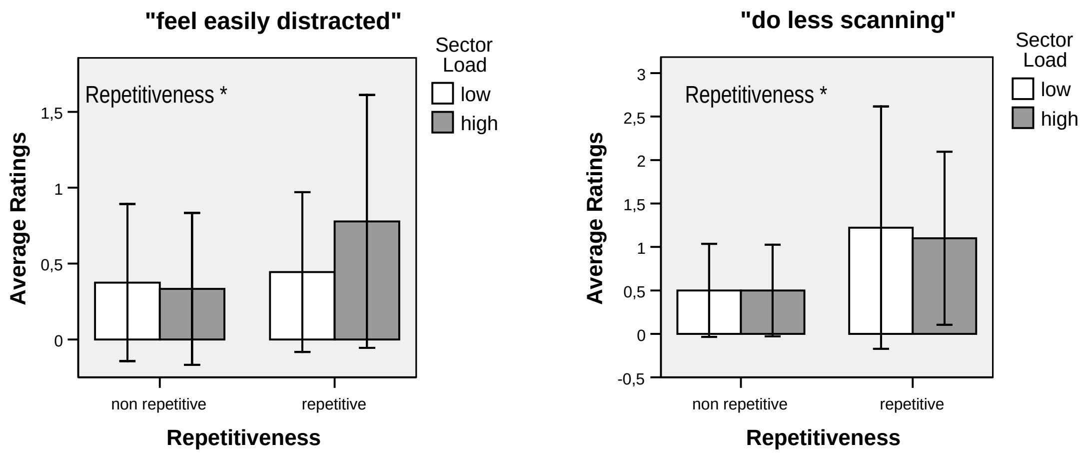


Figure 56: Average ratings and SD on performance-related aspects in repetitive and non-repetitive conditions under increased and reduced traffic load.

The comparison of the different sectors revealed an average traffic load in the compared 90-minute-periods of 50 aircraft per hour in the high load conditions versus 44 aircraft in the low traffic load condition. The distinction between high and low repetitiveness was confirmed by individual ratings of repetitiveness after completion of work periods, however, this effect was not significant (non repetitive: M= 3.21, SD=1.08; repetitive: M= 4.02, SD=1.70; p=.227). Also, traffic density and complexity were rated lower and traffic routine was rated higher in the repetitive high sectors (Table 65; descriptive statistics in Table C-8).

Table 65: Results of Linear Mixed Model Analysis for traffic-related indicators (Study III)

Source	Results ($F_{df\ nominator,\ df\ denominator}$, p-value)				
	Traffic repetitiveness	Traffic routine	Traffic density	Traffic complexity	Traffic difficulty
Repetitiveness	$F_{1,26}$=1.53; p=.227	**$F_{1,26}$=3.25 ; p=.083(*)**	**$F_{1,26}$=11.98; p=.002****	$F_{1,26}$=20.55; p=.000***	**$F_{1,25}$=26.66; p=.000*****
Traffic Load	**$F_{1,26}$=3.36; p=.078(*)**	**$F_{1,25}$=4.19; p=.051(*)**	$F_{1,26}$=1.41; p=.246	$F_{1,26}$=1.44; p=.242	$F_{1,25}$=1.60; p=.217
Repetitiveness x Traffic Load	$F_{1,26}$=.13; p=.721	$F_{1,26}$=.00; p=.972	$F_{1,26}$=.80; p=.378	$F_{1,26}$=.16; p=.691	$F_{1,25}$=.18; p=.678

*Note. 3 missing work periods. ***p<.001. **p<.01. *p<.05. (*)p<0.1.*

There was also no influence of work satisfaction, work experience, and age. In the individual ways to act and morning/eveningness preferences, the sample can be described as rather homogeneous. A table presenting the descriptive statistics for these indicators is included in the appendix (Table C-9).

6.3.4.2. *Effect of different levels of traffic load (Descriptive Hypothesis 4)*

To compare the impact of low, moderate and high traffic load on subjective ratings, the average aircraft count per minute was summarized for 30 minute periods and the ratings of the subjective indicators and traffic characteristics compared (descriptive statistics Table C-10). The cut-off points were based on splits of the total numbers of measurements available for aircraft under control and were 6.4 between the low and moderate group and 8.2 between the moderate and high traffic load group. The Traffic Load Group and the section of the work period were submitted to analysis based on the linear mixed models approach.

The average aircraft count in 30-minute-sections had a significant effect on various subjective ratings (Table 66).

Table 66: Results of Linear Mixed Model Analysis for effects of traffic load on subjective indicators (Study III)

Source	Results ($F_{df\ nominator,\ df\ denominator}$, p-value)					
	Attentiveness	Concentration	Workload	Effort	Feeling Monotony	Boredom
Traffic Load Groups	**$F_{2,110}$=10.93 p=.000*** **	**$F_{2,110}$=9.97 p=.000*** **	**$F_{2,110}$=11.33 p=.000*** **	**$F_{2,110}$=15.23 p=.000*** **	**$F_{2,110}$=5.50 p=.005** **	**$F_{2,107}$=5.62. p=.005** **
Section	$F_{2,110}$=.38 p=.685	$F_{2,110}$=.64 p=.529	$F_{2,110}$=.52 p=.594	$F_{2,110}$=.50 p=.608	$F_{2,110}$=.33 p=.721	$F_{2,107}$=.05. p=.954
Traffic Load Groups x Section	$F_{4,110}$=.91 p=.461	$F_{4,110}$=.73 p=.574	$F_{4,110}$=.42 p=.792	$F_{4,110}$=.66 p=.618	$F_{4,110}$=.36 p=.835	$F_{2,107}$=.30. p=.880

*Note. ***p<.001. **p<.01. *p<.05. (*)p<0.1.*

Further post-hoc analysis revealed significant differences in the comparison of the low and moderate traffic load groups with the high traffic load groups. Controllers reported higher attentiveness, concentration, workload and effort under high traffic load; ratings of monotony and boredom were higher in the low and moderate conditions. Again, no differences in HR (Traffic Load Groups: $F_{2,95}$=.25, p=.782; Section: $F_{2,95}$=.26,p=773, Traffic Load Groups x Section: $F_{4,95}$=.52, p=.723) and HRV (Traffic Load Groups: $F_{2,95}$=.12, p=.891; Section: $F_{2,95}$=.03, p=.968, Traffic Load Groups x Section: $F_{4,95}$=.44, p=.777) were found. Since there was no significant effect of the sections, the following examples in Figure 57 are depicted for the traffic load factor.

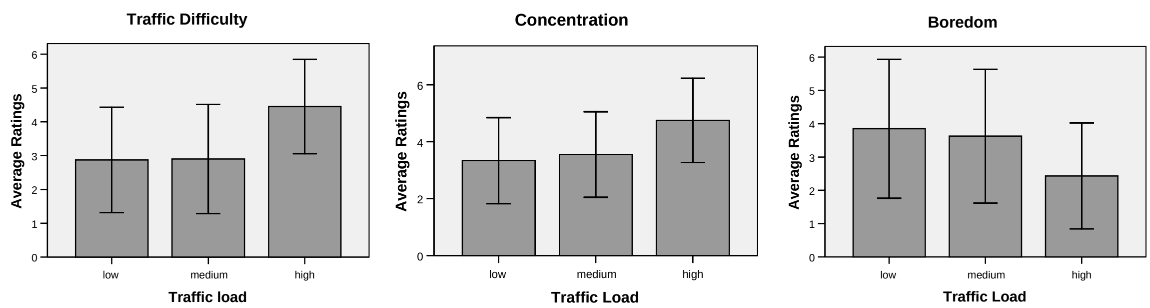


Figure 57: Average ratings of traffic difficulty, concentration, boredom depending on traffic load

Ratings of traffic characteristics revealed that difficulty, density, and complexity were higher in high traffic load, while more routine was experienced under low traffic load. This confirmed the earlier mentioned way to classify the groups in the low, moderate and high traffic load group. No effect was found in the rating of repetitiveness (Table 67).

Table 67: Results of Linear Mixed Model Analysis for effects of traffic load on perceived traffic characteristics (Study III)

Source	Results ($F_{df\ nominator,\ df\ denominator}$, p-value)				
	difficulty	density	complexity	repetitiveness	routine
Traffic Load Groups	**$F_{2,110}$= 13.25 p=.000*****	**$F_{2,110}$= 14.34 p=.000*****	**$F_{2,110}$= 16.18 p=.000*****	$F_{2,110}$= 1.17 p=.313	**$F_{2,110}$= 3.31 p=.040***
Section	$F_{2,110}$=. 43 p=.651	$F_{2,110}$= .34 p=.716	$F_{2,110}$= .86 p=.424	$F_{2,110}$= .02 p=.981	$F_{2,110}$=.15 p=.865
Traffic Load Groups x Section	$F_{4,110}$= .47 p=.760	$F_{4,110}$= .65 p=.626	$F_{4,110}$= .76 p=.553	$F_{4,110}$=1.81 p=.131	$F_{4,110}$= .81 p=.520

*Note. ***p<.001. **p<.01. *p<.05. (*)p<0.1.*

A separate analysis was undertaken in a reduced group for the blood pressure indicators. These participants were only working in the high WEST HIGH sector in the first work period and the EAST LOW sector in the second period. A distinction into low and high traffic load revealed an interaction between traffic load and the work period, depicted in Figure 58.

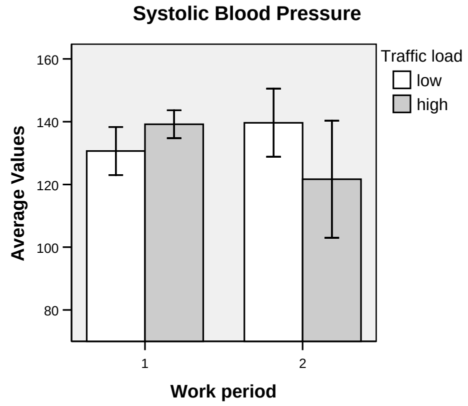


Figure 58: Average values and SD for systolic blood pressure under low and high traffic load for both work periods.

6.3.4.3. The development of mood and critical states during the shift (Descriptive Hypothesis 5)

The SOF scales were submitted to univariate analysis of variance based on the general linear model with study day (i=2) and time-on-shift (i=3) as within-factors. Descriptive statistics are presented in Table C-11 and a summary of the results can be found in Table 68.

Table 68: Results of Analysis of Variance for effects of study day and time-on-shift on subscales assessing mood and critical states (Study III)

Source	Results ($F_{df\ nominator,\ df\ denominator}$, p-value)						
	Stress	Satiation	Monotony	Fatigue	Hedonic Tone	Tense Arousal	Energetic Arousal
Study Day	$F_{1,8}$=.10 p=.765	**$F_{1,8}$=5.92 p=.041***	$F_{1,8}$=.00 p=.964	$F_{1,8}$=.00 p=.962	$F_{1,8}$=2.03 p=.197	$F_{1,8}$=.90 p=.373	**$F_{1,8}$=12.88 p=.009****
Time on shift	$F_{2,16}$=2.03 p=.164	$F_{2,16}$=3.21 p=.067	$F_{2,16}$=2.31 p=.132	**$F_{2,16}$=3.31 p=.063(*)**	$F_{2,16}$=.83 p=.458	$F_{2,16}$=2.19 p=.149	$F_{2,16}$=.48 p=.626
Study Day x Time on shift	$F_{2,16}$=.03 p=.975	***$F_{2,16}$=3.49 p=.055(*)***	$F_{2,16}$=1.01 p=.387	$F_{2,16}$=.38 p=.689	$F_{2,16}$=.86 p=.444	$F_{2,16}$=1.78 p=.204	$F_{2,16}$=2.35 p=.131

*Note. ***p<.001. **p<.01. *p<.05. (*)p<0.1.*

A marginally significant increase in the scores of the fatigue subscale independently of the assigned work periods does however reflect an increase in general fatigue during the work day (Figure 59).

No significant differences were found in satiation and stress; the interaction between study day and time-on-shift was marginally significant for satiation.

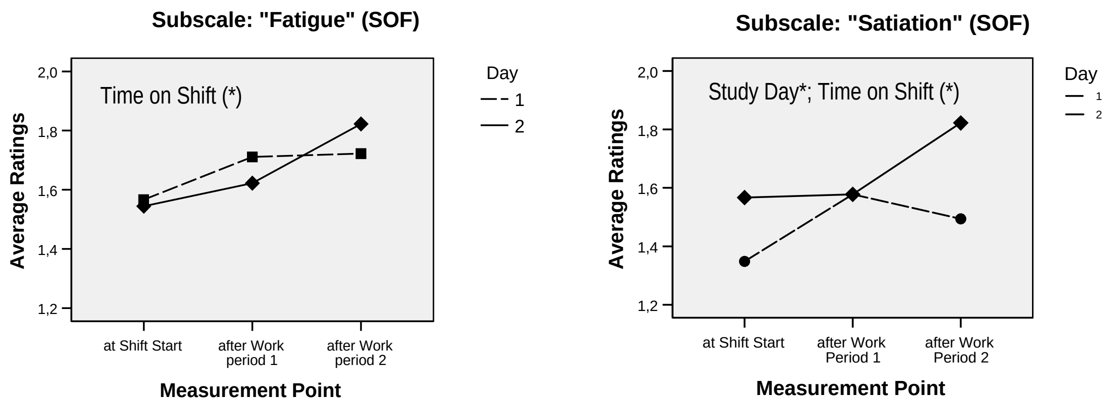


Figure 59: Average ratings of the Scale of Feelings(SOF) subscales for satiation and fatigue in the time course of each study day (Legend: ■ dotted line: Day 1; ◆continous line: Day 2)

In the comparison of the ratings for each study day the scores in the UWIST mood assessment subscale energetical arousal reveal a similar course as in the earlier mentioned subscale satiation.

As a different combination of indicators was collected in the field study, it was not possible to use the same indicators for summarizing a combination of variables as in the simulation study. To assess satiation, frustration (NASA-TLX) and tense arousal (UWIST Mood assessment scale) were combined. Similarly, to reflect fatigue, concentration and energetic arousal were combined. There was no significant effect of traffic load ($F_{1,25}$=.02, p=.901) and repetitiveness ($F_{1,25}$=.81, p=.376) nor an interaction regarding the composed indicator for satiation. In the composed indicator of fatigue, an effect of repetitiveness was found ($F_{1,25}$=.08, p=.786), but no effect of traffic load ($F_{1,25}$=.77, p=.389) or interaction ($F_{1,25}$=1.41, p=.246).

Several indicators were compared for effects of time of shift and day. The development of the subjective indicators attentiveness, concentration, boredom and flow during the work shift revealed no significant effect of Day or Time on Shift. The effect of the work period on strain marginally failed to reach significance. Strain tended to be lower in the second work period ($F_{1,27}$=3.29, p=.081), there was however no difference in the two days ($F_{1,27}$=.14, p=.708) or an interaction ($F_{1,27}$=.28; p=.601). On the other hand, motivation was tendencially higher in the second day ($F_{1,27}$=2.99, p=.095; but significantly lower in the second work period ($F_{1,27}$=8.12, p=.008); there was no significant interaction ($F_{1,27}$=1.75; p=.197). The composed indicator for the state of monotony did not reveal any differences (Day: $F_{1,27}$=.31, p=.584; WP: $F_{1,27}$=1.12, p=.300; Day x WP: $F_{1,27}$=1.47, p=.236).

6.3.4.4. The effects of changes in traffic density (Descriptive Hypothesis 6)

Two different types of analysis were performed to better understand the effect of traffic shifts (descriptive statistics in Table C-12) and traffic anticipation.

Firstly, it was asked how the perceived changes in traffic density were related to the direction on motivation and monotony. The mixed model analysis revealed no effect if perceived traffic load was changing from high to low on either monotony or on motivation. On the other hand, tendencially significant higher motivation and a significantly decreased indicator of the state of monotony were found in conditions where the traffic load shifted from low to high (Table 69).

Table 69: Results of Linear Mixed Model Analysis for effects of subjectively perceived changes in traffic on monotony and motivation

Source	Results ($F_{df\ nominator,\ df\ denominator}$, p-value)	
	Monotony	Motivation
Change low - high	**$F_{1,35}$=5.91, p=.020***	**$F_{1,33}$=4.02, p=.053(*)**
Change high - low	$F_{1,35}$=0.41,p=.528	$F_{1,33}$=.08, p=.777

In a further analysis the impact of the anticipation of the traffic by the controllers was investigated. It was assumed, that if the traffic was predicted as high, the indicator for the state of monotony was lower. For this reason, the predicted traffic ratings before and the ratings after the work period which both concerned the traffic complexity and density were combined in one indicator.

Table 70: Descriptive statistics for expected safety, efficiency, taskload and workload (SET-W) ratings on two days

		M	SD	min	max	N
Day 1	Safety	1.80	.42	1	2	10
	Efficiency	1.30	.48	1	2	10
	Taskload	.30	.68	-1	1	10
	Workload	3.00	.47	2	4	10
Day 2	Safety	1.75	.46	1	2	8
	Efficiency	1.38	.52	1	2	8
	Taskload	.50	.93	-1	2	8
	Workload	3.25	.46	3	4	8

Two groups were formed depending on increasing or decreasing values. Thus, the experience of monotony should be lower in the increasing condition, while elevated in the decreasing condition. This assumption was confirmed by the significant effect of anticipation, where the decreasing condition resulted in higher scores of the indicator of monotony compared to the increasing condition ($F_{1,36}$=11.78, p=.002). A similar indicator was the rating of the expected safety, efficiency, taskload, and workload for the day collected at the beginning of the work shift; descriptive statistics are presented in Table 70.

6.3.4.5. Systematic collection of strategies to mitigate monotony

During the debriefing section the strategies that might help controllers to mitigate monotony were addressed. One of the strategies looked at was the shifted planner exchange. This means that in a total of 90 minute-work periods for both controllers, one of the controllers is exchanged each time with a new controller after 45 minutes. For example, if the EC is scheduled from 14:00-15:30, the PC would be scheduled from 14:45-16:15. To better understand the advantages and disadvantages, the ten participating ATCOs were systematically asked about their experience with that issue. It was mentioned as an advantage by two ATCOs, that during sector hand-over, there is always someone in the sector, who has the picture of the traffic. It depends on the level of the traffic. One person mentioned to prefer to work with the same planner for the whole working period, while two persons stated that they did not mind who of their colleagues was working with them. The exchange on the PC position is however depending on the traffic density at that moment as well as on the perceived competencies of the planner. In high traffic the exchange was perceived as disturbing, and thus requires more time for sector handover. In low traffic density it did not matter; one controller mentioned that a new colleague brings some change. Moderate traffic was perceived as an ideal condition. Concerning the competencies it was mentioned by three controllers that they consider the abilities of the planning controller for the strategy they choose to solve a problem. Thus, the switch to a “good” planner is perceived as positive under high traffic conditions. Further suggestions on strategies how to deal with monotony were classified and summarized in Table 71 under consideration of positive aspects and risks.

Table 71: A collection of strategies to mitigate monotony and related aspects

CATEGORY	ACTION	POSITIVE ASPECTS	RISKS
Non-task related communication	Chatting with colleges	Increase positive mood Helps relax	Avoid subjects that are emotionally involving Keep on scanning the screen Not turn away from radar Awareness focused on the radar
Radar-related activities	Monitor traffic, check everything, where the ac is going, specified level and route (scan screen with 4 ac as with 20 ac) Execute coordination Look at conflicts from greater distance (n=2) Hiding labels (n=1) Give direct routings (n=1)	Remain active	Cannot preplan too much in advance Not very useful, just a game
Sectorization	Flexible sector de-/collapsing by supervisors Flow managers to include the Information of the expected departures and arrivals in the decision	Helps avoid low traffic load	
Mental set	Anything can happen (n=1)		
Additional activities on position	Stand up (n=1) Read (n=2) Computer work (n=1)		You cannot always stand up when you have to control
External activities in break	Exercises Computer-related activities	Variation	Computer-related activities have impact on task

6.3.4.6. The occurrence of personally relevant situations

Personally relevant situations describe any event during the work of the air traffic controller that he or she interprets as having a critical impact on safety. During the measured periods, five situations were rated as personally relevant by the controllers (Appendix C). Two of them occurred in moderate traffic density, one in low traffic density, one in high traffic density and one in complex traffic. One situation was described as strongly deviating from routine, while the others were considered as routine situations. All situations were dealing with unexpected events; none of them was related to technical problems. The events were detected during information collection and monitoring the radar. Four out of the five situations involved team factors and one involved also communication as a factor.

The SET-W ratings were conducted for four of them (Table 72) and revealed that one situation was having a negative impact on safety while the others did not affect safety. The situation mentioned was an unexpected situation related to a military flight that did not have permission to enter the adjacent airspace and was required to hold on a certain point while arranging the situation. Safety was concerned as this holding had an effect on the climbs of other aircraft to reach a certain flight level.

Table 72: Safety, efficiency, taskload and workload (SET-W) ratings for personally relevant situations

ID	S	E	T	W
14	2	1	-1	2
17	2	1	0	3
24	2	1	0	4
25	-1	0	-2	1

Note. n=4. ID=Idendity, S=Safety, E=Efficiency, T=Taskload, W=Workload.

6.3.4.7. The effects of time on shift during night shift

An additional scope was to describe the development of subjective indicators during the early phase of the night shift. Through the focus on one position characterized as repetitive it was evaluated how the indicators would develop during the early phase of the night shift. The scores of the SOF subscales for critical states and the UWIST mood assessment subscales at the beginning of the shift and after the two work periods were submitted to analysis of variance for repeated measures based on the general linear model with the factor time on shift (i=3). Descriptive statistics are presented in the following Table (Table 73).

Table 73: Descriptive statistics for mood and critical states subscales after each work period (WP)

		Stress	Satiation	Fatigue	Monotony	Hedonic Tone	Tense Arousal	Energetic Arousal
Pre-shift	M	1.48	1.52	1.59	1.89	2.58	2.79	2.04
	SD	.38	.44	.41	.36	.21	.32	.29
After WP 1	M	1.47	1.60	1.71	1.93	2.65	2.81	2.00
	SD	.37	.51	.40	.32	.23	.31	.31
After WP 2	M	1.67	1.79	2.09	2.16	2.78	2.79	1.83
	SD	.45	.57	.61	.40	.25	.31	.37

Note. N=9

The results summarized in Table 74 revealed a significant effect in almost all indicators except tense arousal. Also stress did not linearly decrease over time.

Table 74: Results of Analysis of Variance for effects of time on shift on subjective indicators

Source	Results ($F_{df\ nominator,\ df\ denominator}$, p-value)						
	Stress	Satiation	Fatigue	Monotony	Hedonic	Tense	Energy
Time	$F_{2,16}$=3.94, p=.041*	$F_{2,16}$=6.67, p=.008**	$F_{2,16}$=11.69 p=.001**	$F_{2,16}$=4.55, p=.027*	$F_{2,16}$=.512, p=.019*	$F_{2,16}$=.02, p=.981	$F_{2,16}$=.442, p=.030*
Linear trend		$F_{1,8}$=9.21, p=.016*	$F_{1,8}$=13.04, p=.050*	$F_{1,8}$=6.74, p=.032*	$F_{1,8}$=12.25, p=.008**		$F_{1,8}$=10.00, p=.013*

Note. N=9. ***p<.001. **p<.01. *p<.05. (*)p<0.1.

Further univariate tests included time as repeated measures factor (i=2) and were conducted for the ratings of the traffic characteristics and the subjective indicators as well as the performance items. Descriptive statistics are presented and the results of the statistical analysis are summarized in Table 75.

Table 75: Descriptive statistics and results of Analysis of Variance for subjective indicators during each work period (WP) in the night shift

	WP1: M	WP1: SD	WP2:M	WP2: SD	F	p
Traffic density	2.97	1.12	3.37	1.23	1.63	.234
Traffic complexity	3.10	.98	3.33	1.29	.37	556
Traffic repetitiveness	3.60	1.24	4.37	1.17	8.19	.019*
Traffic routine	4.60	1.11	4.70	.99	.11	.745
Traffic difficulty	2.97	1.15	2.77	1.07	1.07	.329
Strain	2.67	1.02	2.93	1.12	.80	.393
Attentiveness	3.40	.99	3.30	1.20	.14	.718
Concentration	3.23	.97	3.17	1.21	.04	.849
Motivation	3.67	.90	2.90	1.63	4.85	*.055(*)*
Feeling of monotony	3.63	1.18	4.33	1.73	5.76	.040*
Sleepiness	2.63	1.16	3.83	1.44	7.55	.023*
Boredom	3.23	1.29	3.97	1.53	8.13	.019*
Mental demand	3.23	.94	2.67	.89	4.06	*.075(*)*
Temporal demand	2.27	1.18	2.33	1.25	.06	.811
Effort	2.80	.96	2.87	1.12	.03	.861
Performance	4.97	1.40	4.83	1.19	.27	.613
Frustration	1.90	1.14	2.00	1.05	.10	.761
Flow	1.40	1.37	1.10	1.51	2.25	.168

Note. N=10; df=1, dferror=9. ***p<.001. **p<.01. *p<.05. (*)p<0.1.

Table 76 presents the descriptive statistics and results of the statistical analysis comparing behavioral indicators using the same procedure.

Table 76: Descriptive statistics and results of Analysis of Variance for behavioral indicators during each work period (WP) in the night shift

	WP1: M	WP1: SD	WP2: M	WP2: SD	F	p
not feel focused anymore	.11	.33	.56	.88	1.73	.225
ask to repeat call from a/c	.89	1.05	.78	.67	.31	.594
not understand R/T	.44	.73	.44	.73	.00	1.000
spot a conflict only 1-6 minutes before	.22	.67	.22	.67	-	-
feel like working less precise	.33	.71	.78	.97	2.29	.169
make small mistakes (e.g., input errors)	.56	.53	1.00	.71	6.40	.035*
feel like getting behind in work	.00	.00	.11	.33	1.00	.347
feel like doing less pre-planning	.22	.44	.67	.71	6.40	.035*
work slower as usual	.56	1.01	.78	.83	1.00	.347
not knowing a/c on frequency	.00	.00	.11	.33	1.00	.347
forgetting routine co-ordination	.00	.00	.11	.33	1.00	.347
react slower as usual	.33	.50	.78	.67	6.40	.035*
pay less attention to detail	.44	.73	.67	.87	2.29	.169
surprised by call	.33	.71	.22	.44	1.00	.347
miss a call	.44	.53	.56	.53	1.00	.347
feel easily distracted	.44	.53	.44	.73	.00	1.000
look for traffic that calls in	.78	.67	.56	.73	1.00	.347
overlooking obvious problems	.22	.44	.33	.50	.31	.594
do less scanning	.89	.78	1.00	1.00	.08	.782
longer pause initial call pilot & identification	.67	.71	.89	.60	1.00	.347

Note. N=9; df=1, dferror=8. ***p<.001. **p<.01. *p<.05. (*)p<0.1.

6.4. DISCUSSION

6.4.1. The Effects of Task Characteristics

Two hypotheses addressed the effects of repetitiveness and traffic load on monotony. The results supported the alternative hypothesis concerning the main effect of repetitiveness. There was no effect of traffic load.

Monotony has been assessed with a composed indicator deducted from previous studies which contained inverted heart rate, subjective ratings of sleepiness and feeling of monotony. It was confirmed that monotony is not only a consequence of situations characterized by repetitiveness and homogeneity, but also of low traffic load as a reinforcing factor in the repetitive sector.

In summary, it was found that repetitive aspects in the work of air traffic controllers are well reflected in the subjective perception of the situation. It is not the reduced workload and strain, but the impaired attention, concentration, and motivation which do require further initiatives to mitigate these effects.

The overall indicator for 90-minute-periods comparing high and low traffic load was not found sufficient to confirm a general difference between different traffic load effects, but a more detailed analysis of different traffic load levels based on the actual traffic count demonstrated that monotony was also higher in low traffic load. In addition, a look at further indicators showed that low traffic conditions caused significantly lower levels of concentration, attentiveness, and higher boredom. The comparison of low, moderate and high traffic load groups revealed that an average minute-by-minute count of less than eight aircraft for a longer period of time (30 minutes) has the potential to result in difficulties concerning concentration and attentiveness as well as increased boredom and feelings of monotony in the investigated sectors.

Overall, this confirmed the theoretical assumption that repetitive traffic situations which are moderately difficult and have a constant traffic load need to be as well considered for the analysis of monotony as the well-known very low traffic situations.

An unexpected finding was that the effect of repetitiveness on HR, which was found in both simulator studies, was not present in the field study. There was however a difference depending on the traffic load. This can be explained with the tendencies of the controllers to self-activate themselves even in lowly demanding working periods. Activities observed were talking with planner controllers as well as with controllers in neighbouring sectors. Despite the positive side-effect of keeping the physiological activation level high, situations were noted where unwanted impacts concerning the distractions by external tasks were observed when returning to the radar.

However, the analysis of individual reaction patterns turned out to show interesting results. One example describes the occurrence of an extraordinary situation under the condition of low traffic load/non repetitive traffic, which resulted in an immediate increase of heart rate during the period in which the problematic situation was solved (Figure 60). As it can be seen in the reaction pattern for heart rate, a relevant event occurred between 15:15 and 15:20 (UTC) which increased the heart rate significantly. However, within 10 minutes the situation was successfully solved and thus the heart rate decreased immediately to the level before the event. This might be discussed as successful immediate recovery under ideal conditions where the controller did not experience high workload during the work period.

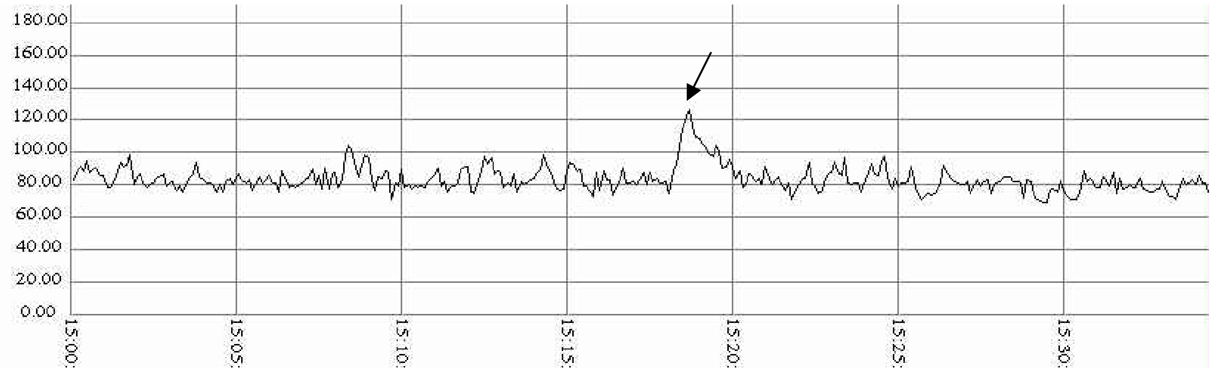


Figure 60: Development of the heart rate in the work period of a controller who experienced a very rare out-of-routine situation (marked).

A marginally significant effect of time was found in systolic blood pressure in the afternoon period. Although measurements took place at the lower EAST sector, which is rather complex because of required sequencing of inbound traffic to Budapest, the overall blood pressure decreased during the work on this position (Figure 61). It is remarked, that participants whose lower and higher sectors were collapsed in a combined EAST sector, were excluded from this type of comparison to maintain similar conditions. The non significant results in blood pressure indicators during the WP needs to be interpreted in consideration of the very low number of available measures (n = 8).

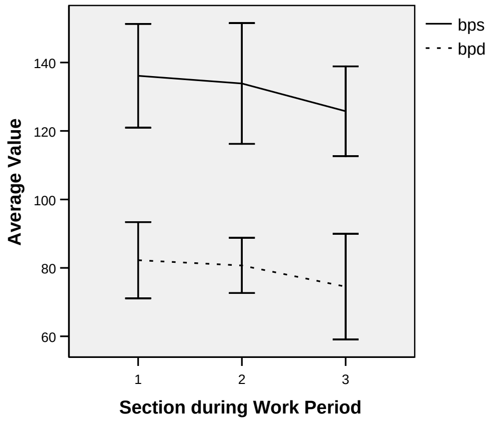


Figure 61: Development of the systolic (bps) and diastolic (bpd) blood pressure during the afternoon work period in the lower EAST sector.

Again, the analysis of individual cases turned out to be successful to reflect effects on blood pressure. Such a situation was the occurrence of a critical situation, where control techniques were not ideally deployed. The following interview revealed a strong emotional involvement of the concerned controller and highly increased blood pressure was also subjectively attributed to this situation. Figure 62 demonstrates the occurrence of the event and the increase of the systolic blood pressure (comparable for diastolic blood pressure) around one hour after the event in the last baseline measurement after work period 2. This development presents a type of sleeper effect; it is remarkable that during the rest of the work at the position the controller's physiological activation remained rather low. Also, the same effect was not reflected in heart rate indicators.

In Figure 63 an example of a different controller is presented who experienced highly increased blood pressure during the first section of a work period characterized by high traffic load and also subjectively rated as such. In this case the effect was also reflected in increased heart rate. Finally, blood pressure was also increased in the measurement following the work at position.

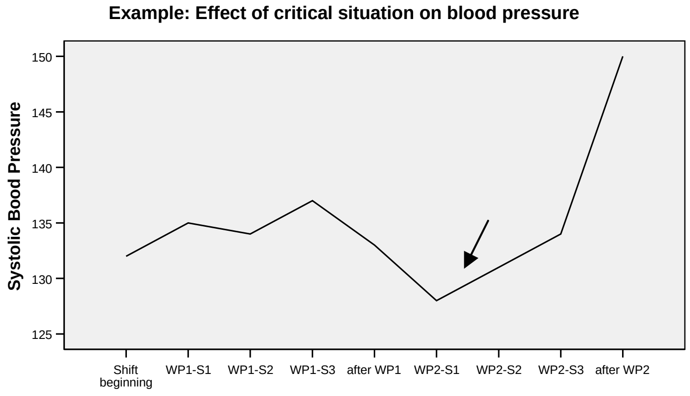


Figure 62: Example of the effects of the occurrence of a critical situation on systolic blood pressure

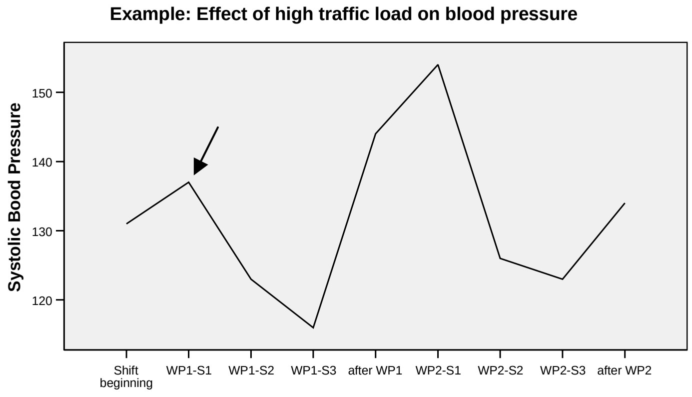


Figure 63: Example of the effects of high traffic load on systolic blood pressure

Certain influences of changes in the work environment need to be mentioned which reflected the emotional involvement of air traffic controllers in work-related issues and relates to ratings of work satisfaction. The before mentioned sleeper effect is very well known in stress research, but not investigated for short-term strain, and could not systematically be considered in the current study. However, the increased blood pressure in emotionally involving situations might be discussed in relation with the general need of critical incident stress management techniques in ATC.

It is also discussed that the results may well relate to the sector complexity, if the structure of the TOP and LOW sectors at the centre is considered. However, complexity and variety are independent factors, as it was described by Zapf (1995). While complexity reflects the number of sub-goals necessary to complete a task, variety refers to the actually visible actions undertaken to complete the task. Thus, this is a relevant argument to assess the effects of repetitiveness as causing variety independent of sector complexity, as it is related to the necessary sub-goals when executing the task.

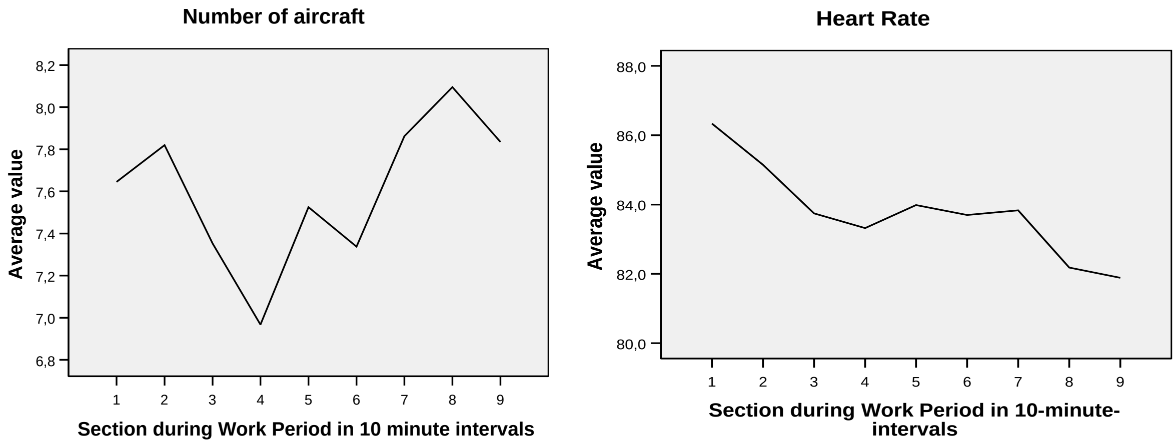


Figure 64: Average heart rate and number of aircraft in 10-minute-sections during all work periods

Unexpectedly, if low, moderate and high number of aircraft were compared only for the repetitive sector, no difference was found in the indicator for the state of monotony, physiological activation or sleepiness or boredom. The effect was however related to workload, effort, concentration and attentiveness. This indicates that uneventful conditions do not directly relate to increased monotony as long as there is some traffic to deal with. It can however be argued that even in the low condition, as it was undertaken with median-split, a certain variety of aircraft count was contained.

The development of the indicators that were collected in the worked periods during the early part of the night shift additionally supported the results of the day shift as the second period was rated higher for repetitiveness, there was however no difference in traffic density, complexity, and routine. At the same time, motivation decreased, and the subjective feeling of monotony, sleepiness and boredom were elevated. Due to the increased ratings in the fatigue-subscale and the reduced ratings in energetical arousal, it is however not possible to neglect the impact of fatigue in the interpretation of the ratings. Even though, attentiveness and concentration were not reduced which does not further support the assumption of a stronger influence of fatigue.

Finally, an overall comparison of heart rate and the number of aircraft under control in 10-minute-sections during the work periods illustrated in Figure 64 revealed that independent on the increased traffic demand towards the end of the work periods the heart rate continued to decrease and indicated suboptimal activation towards the end of the work on position. This did not differ in any of the work periods. It is remarkable, that the number of aircraft under control follows a cubic trend, while heart rate linearly decreases. Thus, the demonstrated relationship of the number of aircraft and heart rate (e.g., Brookings et al., 1996) would require further investigation with reference to the time-on-task factor.

In consequence, the described operator state may be suboptimal for work execution as adaptation to changing situations may not be successful and thus impose a risk factor for maintaining optimal performance. This was reflected in a different course of the task demand imposed by the number of aircraft under control and the heart rate indicator.

6.4.2. The Effects of Influencing Variables

The hypothesis that individual factors influence the outcome of monotony was supported by the indicator of the initial state of recovery but not by the data of the initial strain state and boredom proneness. For example, reduced levels of the initial state of recovery at the beginning of the shift had a favorable impact on monotony. Several variables that were expected to have an impact on monotony were not considered in systematic statistical analysis, as the sample was homogeneous.

An additional indicator of interest was the relation between initial expectations before the work at position as well as the initial ratings for expected safety-workload-efficiency. It was expected, that if there was congruence in the predictions with the actual situation, the controllers would be able to adjust their mental set better. Anticipation is one of the core processes, such as planning (Kallus et al., 1998, p.49), and decisions are undertaken according to an ad hoc criterion, which is strongly affected by habits or situational or other factors the controller is not aware. Also in the study of Rau and Richter (1996, p. 275), the anticipated task difficulty had a greater influence on psychophysiological strain than the subsequent evaluation. In the own study, deviations from the anticipation of conflicts and traffic load in a WP influenced the development of a state of monotony.

The results give rise to the idea that a controller can well cope with either high or low traffic load extremes if there are no initial suboptimal states at the beginning of the shift. However, the long-term effect and the accumulation over time of fatigue or other suboptimal states is not well understood yet.

6.4.3. Further Evidence for Different Strategies

A systematic consideration in the study design concerned the effects of shifting changes in traffic load. As already seen in the simulation, an increase in traffic within certain limits resulted in higher motivation and was confirmed also in the field settings. On the other hand, a perceived decrease in actual traffic load had no effect. This could be supported with one statement of a controller that “one swings with the traffic, when the traffic goes down”. As well it might be explained that the changes in traffic load were not extreme, and thus did not lead to exhaustion.

Several strategies have been determined and repeatedly mentioned by the controllers to support the maintenance of optimal psychophysiological conditions. The shifted planner exchange was especially considered and pros and contras defined. Traffic load and the competence of the planner were seen as crucial factors for a successful exchange.

Concerning additional strategies to mitigate monotony some variety between the teams was noted, but in general the suggestions addressed balanced communication with colleagues, work-related activities during the work at position, and more flexible sectorization as well as alternative activities in the breaks. However, as especially the communication aspect may have critical side-effects, this strategy needs to be carefully deployed.

Overall, the applied strategies seem to tackle two main scopes. The first one seems to be related with maintaining a positive state and a good mood. The second one appears to be related to the activation aspect. The importance of both aspects was very well known to provide good performance by the interviewed controllers, they were however not well enough considered in any suggestions. The importance of mood for performance has been confirmed by different studies (e.g., Gendolla & Krueksen, 2001). Also the relevance of the state of activation has been confirmed, despite frequent criticism as summarized in Muse, Harris, and Field (2003). The controllers seem to be well aware of the potential negative effects of strategies like chatting, still, critical situations happen. For example, one situation was observed, in which difficulties to turn back to the radar were noted after intense involvement in chatting.

Still, how to deal with boredom or monotony is not a systematic aspect considered in the training and therefore any strategies especially taught to trainees remained scarce. Moreover, at the initial state a trainee might be less prone to experience monotony, even though the issue of boredom has already been mentioned in the EUROCONTROL Common Core Guidelines for Initial Training (2004). Overall, a repertoire of strategies to apply in different situations of monotony still should be developed that clearly include the advantages and disadvantages of each. An individual controller should have the opportunity to find out for him- or herself, which one is better suitable.

6.4.4. Are there Critical States in the Field?

A similar procedure to investigate critical states was applied in the field study. That no significant differences in the SOF-subscales were found is partly explained with the fact that the questionnaire was provided after the work period. It is possible that the effect of a small rest break before filling in the questionnaires influenced the ratings positively.

On the other hand, considering the development over time a different picture emerged in the satiation indicator. It is a remarkable result, that in comparison of the first and the second day independent of the assigned work periods a significant increase in the subscale satiation occurred. This might have been influenced by ongoing changes in the work organization at the control centre, which was also reflected in the correlations between the satiation and job satisfaction. In addition, controllers received work-related information that was intensely discussed in the ops room environment before the second study day.

An additional analysis used the indicators reported by Richter et al. (2002). Because of a different focus, different indicators were used. The composed items for satiation and fatigue did however not reveal any significant differences in the course of the day. Thus, the results did not contribute to further enlighten the distinction between critical states.

6.4.5. Methodological Issues

The following subsection discusses weaknesses of the present study, which are to be considered in the interpretations.

That the number of aircraft may not necessarily relate to increasing heart rate, as it was frequently demonstrated in other studies, and that cardiovascular predictors developed differently as expected, shows the relevance of carefully selecting the procedures and intervals of the measurement as well as any other occurrences in the operational environment. For example, in line with Backs (1998) research, a different development of HR and HRV is possible depending on their coupling mechanisms. Ten-minute-intervals were considered as sufficient to reflect dynamic changes and adaptation mechanisms. The collection of special events and perceived traffic characteristics was necessary to detect related changes in the physiological measures.

It was feared that the study leader's presence in the observations might have had an impact on the controllers behaviors especially in low traffic density conditions and thus different avoiding behaviors would occur. As stated, once ATCOs were involved in their task, they did not think any more about the observer's presence and rather interpreted it similar to an instructor's role. All controllers perceived the collection of heart rate measures rather positive and the ECG recorder was not perceived as intrusive. Two participants perceived the automatic blood pressure monitoring as too obtrusive during the work at the position, and thus did not accept to keep it after an initial testing phase. It was perceived as disturbing, since these participants were using mainly the left hand for frequent activities at the radar and thus felt restricted in their activities. Especially the measurement of blood pressure as well as the ratings of performance required trust in the experimental leader.

Some items of the questionnaires and scales were perceived as difficult. For this reason, an additional list of translated items was provided for the mood scale. To conduct extensive interviews in a foreign language was successful, it was however imposing greater stress for some participants, which might be reflected in some of the higher blood pressure values automatically measured during the rest breaks.

Still, the selected sample needs to be discussed concerning a specific interest in creating variety in the work environment, as reflected in the involvement in further functions in the job. However, because of the importance of subjective ratings and interpretations, a non-interfering method to describe its course should be preferably applied during work on position to better understand the course.

6.5. SUMMARY

The objective of the field approach was to investigate the factors of the simulation studies in real work-settings. The complete within-subjects design contained the factors repetitiveness and traffic density. Additional independent variables of interest were experienced changes in traffic load from high to low or low to high density. Dependent variables contained heart rate and heart rate variability, systolic and diastolic blood pressure, traffic-related, subjective and performance-related ratings similar to the ones applied in the simulation environment, and individual characteristics such as initial states were collected as moderator variables. Ten air traffic controllers with an average age of 36 years were measured in four counterbalanced work-periods of 90 minutes on two study days. The results showed that sectors characterized by increased repetitiveness in the required actions to complete the work did not only lead to positive effects such as reduced workload and strain. Controllers also experienced reduced motivation, attentiveness, concentration and increased boredom. Some of these effects were even more pronounced under reduced traffic load. These effects were however not reflected in physiological measures analyzed in the overall condition. Nonetheless, the description of individual cases showed covered physiological effects which turned out to be rather the consequence of clearly distinguishable occurrences on the individual level. Delayed and immediate effects on blood pressure were observed under consideration of personally relevant occurrences.

On an individual level, the initial state of recovery was confirmed as a moderator variable that influenced the development of critical states. At the same time the collection of organizational processes helped to understand part of the results. Potential mitigation strategies were determined in debriefings and comprised shifted planner exchange, balanced communication, flexible sectorization, as well as alternative rest break activities.

Again, the study demonstrated the importance of a multivariate approach combined with observations in the field to explain unexpected results.

7. GENERAL DISCUSSION AND CONCLUSIONS

7.1. OVERVIEW OF RESEARCH OUTCOMES

The research described here was dedicated to a systematic investigation of factors that evoke or contribute to a state of monotony in ATC. Several relevant issues were identified in the review of the literature and subsequently addressed in experimental studies. Because the results of previous studies could not directly be applied for ATC, a focus of research on monotony was needed for this specific environment. This was even more relevant, as the role of monotony cannot be ignored as far as it concerns the actual situation in the operational environments as well as the ongoing development of future concepts. Thus, the general aim of these studies was not only to define relevant factors for monotony and to describe its appearance, but also to define actions that could be undertaken to mitigate the effect of monotony.

Research questions were defined around these three objectives and were aimed at:

- determining factors evoking and influencing monotony,
- the description of monotony and distinction of related concepts; and
- the collection of strategies for mitigation or countermeasures.

Concerning the research questions around the first objective the task factors repetitiveness and traffic load/traffic density were found to have a significant effect on a state of monotony. This was true in simulation settings as well as in field settings. It needs to be considered, that the manipulation of the experimental variables in the field was subject to slightly different criteria. Dynamic Density was varied in the laboratory while traffic load was manipulated in the field. Both variables were however expressing different forms of (un-)eventfulness.

An additional factor influencing the results on monotony was the initial state of recovery. Boredom proneness was not confirmed as a significant factor, the results suggest however further investigation. The experience of flow counteracts monotony in both lab and field settings.

Further effects related to time-on-task were not surprising and confirmed the relevance of the time effect. This leads into the discussion of the research questions related to the second objectives that is the description of states of monotony and the distinction of critical states. As was found out, critical states develop differently over time. This needs to be discussed separately for the simulation and the operational environment. In the simulation the traffic effect was strongly affecting the first hour, and fatigue effects during the second hour led to mask the effects of monotony. In the operational environment, the appearance was slightly different as the effect of repetitiveness was analyzed for the aggregated working period, since it was not possible to collect ratings during the work at position. Even though the task characteristics did have the expected impact on monotony, other critical states need to be regarded, as they are influenced by cognitive interpretations. Fatigue was stronger than monotony after a certain time on task in the simulation, while in the operational environment different factors were important. There, organizational influences affected satiation.

The last objective addressed the mitigation strategies and asked for the effectiveness of predefined and easy to implement countermeasures. One such countermeasure was physical break activity that reduced the effect of the task on the state of monotony. Additional countermeasures came up during the studies. The effect of traffic changes from low to high was positively perceived in the simulation, but the effect was not found again in the field. In the field, a shift from high to low resulted in negative states. As initially addressed, countermeasures may effect three different levels, namely the task, the individual and the organization.

In conclusion, the study has shown that low traffic and repetitive conditions do have undesirable effects on the subjective state of controllers and thus might help to explain the occurrence of critical situations. On one hand the importance of routine for an air traffic controller is well known, since it enables to deal with a high amount of requirements imposed by the traffic. On the other hand the same situation implies the risk of monotony. Also, as routine predominantly consists of habitual processes in task execution, it may contain the danger to lead to complacency or over-confidence contained in the feeling that something runs smoothly and thus less effort is invested. This was also confirmed by interview statements. Interestingly, the participants also felt more confident in their work when they had more to do. This has immediate impacts not only for the operational environment but also for concept development.

7.2. DISCUSSION OF THE RESEARCH APPROACH

Within the defined framework three studies were conducted to answer the range of research questions. Subsequent to a small-scale study that was arranged to verify the experimental set-up, the relevant factors were included in a main study to investigate their significance in the controlled setting of a laboratory study. After determining a prior model, the results were evaluated in field conditions.

While the simulation studies manipulated the factors and controlled for all influences, in the work-setting approach perceived variety was additionally included as a moderator variable. The disadvantage was however that the traffic was difficult to predict and thus it was relied on the sector characteristics and experiences of the experts to determine the experimental conditions.

As previously defined, a multivariate approach was necessary to show the effects, since different levels of measurements are each connected with disadvantages and advantages. This approach was successful to find common factors in the lab and the field and revealed also where contradictions do occur. The results of the two traffic factors were repeated for all conditions.

The application of quantitative and qualitative methods further brought up additional information. Reconstruction interview techniques and debriefings supported the quantitative results and contributed to gain additional insights.

Else, the quality of the measurement instruments needs to be discussed. Throughout the studies, SOF was not very effective in reflecting group differences. Also, the mood subscales were not showing the expected pattern. The following reasons help to explain this: the SOF scale was perceived as long and the transfer of some items to the own working environment was difficult. Also the background of the SOF does not support a distinction of different states, as it was developed with stimulus- and response scaling procedures and not with factor analysis. The mood assessment might have been especially influenced by the measurement after the task had already ended.

Ideally, the same participants would have participated in the lab and the field. This was strived for, but could not be realized. Even though, the selection of the samples, which consisted of experienced air traffic controllers in both situations, counteracted this limitation. To compare the similarity of the overall reactions in both settings, the absolute HR change in relation to the baseline is illustrated in Figure 65. Even though the range of the difference is comparable in both settings, it is remarkable that the HR average of the operational sample is markedly higher.

Final discussions concern the generalizability of the findings. The findings regarding the effect of task characteristics on the subjective level can be generally applied, as they have been confirmed in different settings and with independent samples. Because of differences in the work organization of the centers, some of the results of the field study are specific for the local environment. Else, that the physiological indicator did not reveal any differences was already argued for.

A comparable observation was also made in a study reported in Kohlisch, Kuhmann, and Boucsein (1991) who objected to validate their findings on the effects of varied system response times in field settings. Partly different results were explained with different requirements in the real world and allowed them to advance their understanding of the research problem.

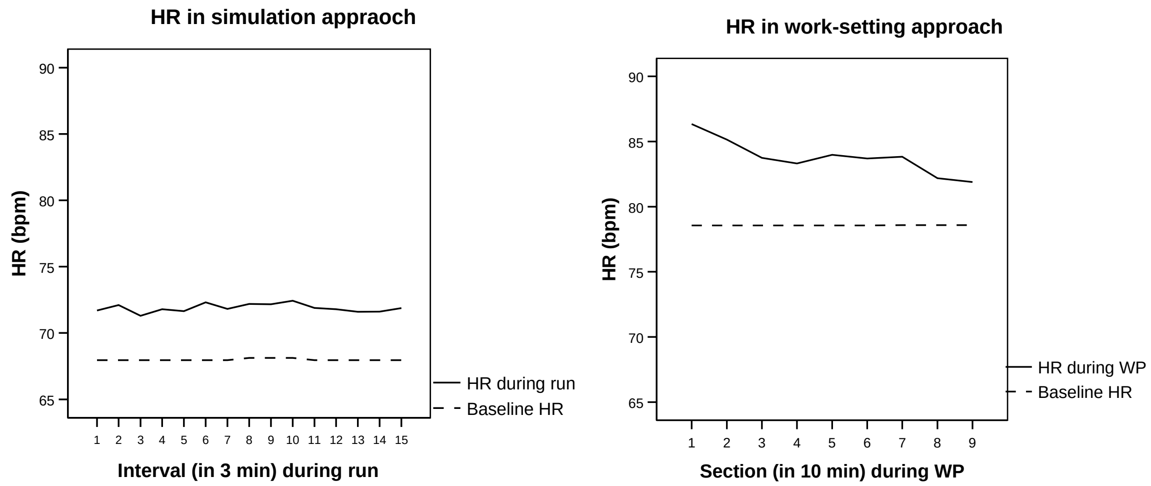


Figure 65: The comparison of the development of uncorrected heart rate in relation to baseline measures in simulated and operational settings

7.3. DISCUSSION OF THE THEORY OF MONOTONY AND RELATED CONCEPTS

This chapter discusses the results with a special view on the theory of monotony and alternative explications and includes also the critical states.

The theory presenting the background of this research was proposed by Bartenwerfer and assumed that monotony occurs in simple tasks of a repetitive nature or low stimulation. The resulting state is characterized by physiological deactivation, sleepiness, boredom and connected with performance fluctuations. In the research activities, the requirements were considered in the manipulated task demands in the lab and field studies. A composed indicator from the strongest measures representing each level clearly supported the assumption that monotony emerges as a result of certain task characteristics. A more fine-grained analysis showed however, that the results were not as clear as expected when analysing the physiological indicators and even contradicted the assumptions. There are several possible explanations for this. It might be possible that more effort is invested to maintain an appropriate functional state in the ops room than in the simulation environment. Controllers generally know that appropriate physical activation is highly important. Thus, the operational experts do have a different way of activating themselves, which might be less efficient as the one used by active controllers, who are automatically reacting to a certain situation. This seems to fit in the pattern to explain monotony as habituation, as in the lab study also SCL revealed a similar course. Habituation might be a potential explanation for laboratory conditions; however, it is not appropriate for the field. In the field HR did not reflect a significant effect of different task conditions, the effect of HR was rather visible in individual reactions to increased task demands.

Consequently, the results need to be seen differently in the simulation and field environments. Increased HRV in the repetitive condition is consistent with the finding of increased HRV in the high boredom/monotony group of Thackray et al. (1975). Perkins and Hill (1985) explained HRV increases under boredom with the fact that an individual may reject demands and cease to process task relevant information or process it intermittently. At the same time, they found that on a cognitive level individuals constructed a situation as less differentiated and more homogeneous.

Also it could be argued that a longer and less attention concentrating task might lead to a rejection of task demands to induce a lower mental load. The malleable resource theory (MART) explained increased errors in underloading situations in terms of adaptation of the state to task demands. This meets the results found in the simulation, but not in the field. Finally, this model does not include any flexible regulation mechanisms, as does the state regulation model of Veltman (2004), which would also allow explaining the dissociation in the subjective and physiological measures.

Another concurrent assumption was that underloading tasks would result in higher workload, as was found in numerous vigilance studies (e.g., Szalma et al., 2004). In that kind of research, workload is usually increased with ongoing monitoring time, whereas in the current research the workload was rated lower in the repetitive conditions. At the same time, also concentration and attentiveness were lower and monotony higher, which is again in line with vigilance studies. More than that, workload did not generally increase with time-on-task and was only affected by higher dynamic density. Vigilance studies also reported increased boredom, which was partly found in the current study, but did not show an inverted course compared to workload. This tendency in workload ratings may be caused by the special nature of the ATC environment, where controllers' frequently associate workload with the amount of aircraft in the sector or with the complexity of the sector, as aircraft count is a common predictor in use to support decisions in sector management. This supports a distinct consideration of the concept of monotony as a consequence of repetitive tasks that require activities compared to vigilance operationalized in monitoring tasks.

Another frequently applied explanation related to underload was based on the stress concept. One of the basic problems related to use stress as an explanation is the different understanding of stress as pointed out in Chapter 2. After the first small-scale study, this argument can however not be totally ignored and would need further investigation, if the long low density condition really was showing effects close to stress phenomena. Stress in terms of an increased excretion of stress hormones was observed in passive monitoring (Johansson et al., 1996). To totally exclude the stress hypothesis for low traffic load of long duration, a further systematic investigation of this aspect would be necessary. It is however hypothesized, that such individual presets might be rather interpretable in terms of satiation with increased irritation and tension. It might not be perceived as a threat unless a personally relevant experience would have already conditioned an ATCO to regard such a situation as dangerous. The current simulation data support increased satiation independent on the experimental manipulation.

Some further aspects that came up in the course of the study concern the relevance of expectations and routine. The expectation is a very important aspect of monotony. If someone predicts something to be monotonous, someone might adapt his or her state to the situation already in the expectation of less effort needed. On the other hand, if someone is aware of the danger of developing a critical state, one might already develop a strategy to counteract and thus less likely develop the perception or expectation of something being monotonous.

The state of monotony is seen as a consequence of exposure to situations that may be either of an uneventful or repetitive nature. However, it needs to be considered distinct from routine. ATC is guided by hierarchically organized anticipation-action-comparison loops that are even true for highly automated habits which reoccur in every shift (Kallus et al., 1998, p. 26). On one hand the importance of routine for an air traffic controller is well known, as it enables to deal with a high amount of task requirements, on the other hand it implies the risk of a monotony state. Also controllers do have a heterogeneous understanding of monotony. It is connected to situations when there is no special requirement by the traffic, but generally distinguished from boredom that is rather occurring in low traffic situations. Routine is connected to experience, and most of the time perceived positively. But, as routine predominantly consists of automatic processes in task execution, it contributes to complacency or over-confidence contained in the feeling that something runs smoothly and thus less effort is invested. In consequence, such a state may be suboptimal for work execution as adaptation to changing situations may not be successful and thus impose a risk factor for maintaining optimal performance.

One component of routine is habits, which ease an efficient work style. But they do have also the potential to contribute to operational errors, if not correctly recognized. Xing and Bailey (2005) defined different categories for operational errors that exceed the cognitive capacity from analyzing runway excursions and defined habit interference as one of them. “This means that if a controller receives a piece of information that is in conflict with what he/she had learned by experience, unless he/she makes an effort to suppress the response from the experience, the previous experience can lead to misinterpretation of the sensory input.” (p. 654). Even though this description is focused on cognitive processes, especially related to perception, in a more general form it can be applied to explain the longer conflict resolution in the repetitive condition of the simulated ATC study. This category of habit interference is also similar to the set effect reported by Luchins (1942), but until today not systematically investigated in ATC. However, with the focus on cognitive processes it does not contain the consideration of the activity to solve the conflict situation.

Finally, the models that were used to explain workload can also be seen in relation to the distinction of critical states. It was seen that the reaction pattern is very complex and not simply comprisable in one workload indicator. A pattern of satiation seems to be related to monotony and fatigue. Also job satisfaction plays a role in the ops room, even though a systematic interaction with critical states could not be determined due to the small sample size.

Alternatively, it might be discussed if monotony and fatigue are distinguished operator states. Monotony appears to be an independent state as long as time-on-task had a minor impact. As soon as this effect gets stronger, any influences of traffic characteristics diminish and thus the fatigue effects remain. Hockey (2003) described the adaptation of strategies to cope with this fatigue.

The assumption of independence or dependence did not result in any satisfying conclusions. It depends on the indicators that are used, which approach is preferred. However, the approach is favored to describe monotony as a task-based state and fatigue as an energetically based state. With relation to the earlier discussed stress hypothesis, this concept would be rather seen as skill-based, as one would be rather stressed by a perceived threat when resources are lacking. In recent work satiation was seen as a motivation-based state (Schultz-Hardt et al., 2001). This can however not be supported by the data in neither the simulation nor in the study in the operational environment, as motivation and irritation were also partly affected by the varied task characteristics.

Either scale used to distinguish critical states has the scope to define the most suitable countermeasure, what needs to remain focused as a primary objective for work environments. So far the assessment of multiple levels was superior to simple workload measures, despite the currently unresolved issue of typical measures for all critical states.

7.4. PROPOSAL OF A MODEL OF MONOTONY IN ATC

Based on the results obtained in the research activities, the scope of this section is to discuss an adapted model[13] of monotony for the field of ATC. The results provide support for a model of monotony that distinguishes uneventful and repetitive work conditions as the basis for a state of monotony. This model helps to understand the phenomena that do occur around the state of monotony and to define at which point the countermeasures need to act upon.

Figure 66 presents the components of the model which supports any decision for a systematic consideration of monotony within the air traffic control setting.

The model basically contains two types of task characteristics relevant for monotony, as it was discussed and confirmed in the research activities. Both need to be considered because of the variation in the task. The difficulty is a further precondition. Other performance-shaping factors on the individual and organizational level contribute to the development of a state of monotony.

The distinction of different potential critical states during the execution of tasks, which are characterized by these two types, is suggested. The dominant pattern of each state can only be described if multiple measures are applied, as it also allows explaining an eventual dissociation between measures.

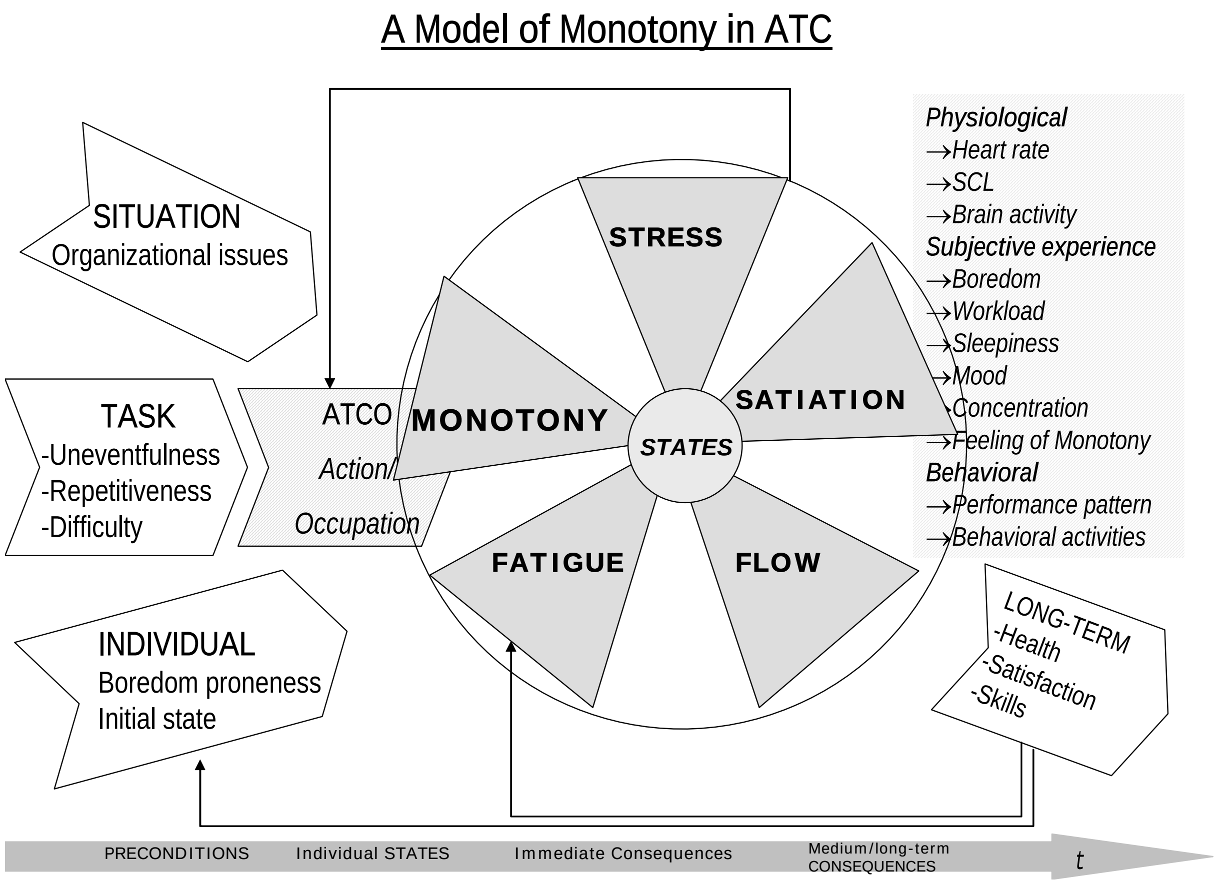


Figure 66: Model of Monotony for ATC

[13] Different definitions for models are currently in use in the scientific environment. In the current work, a model is defined as proposed by Stachowiak (1973) who characterized a model by the representation of selected features of the environment. At the same time, the components represent a restricted selection and depend on the person who created the model with regard to its usefulness.

A state of monotony is supported by an increase in a composed indicator of subjective sleepiness and feeling of monotony. The third component, which is a cardiovascular parameter, does not indicate deactivation in both simulation and field settings as expected. This can be explained by alternative activities. Thus, to understand monotony in field settings, it might not be sufficient to record physiological measures, but also behaviors need to be analyzed. An increased observation of communication outside the task might indicate monotony in the task.

The wheel of fortune analogy contains the idea that many factors influence the outcome, which is not systematically predictable, because small changes in the context lead to a different form of action and occupation during task execution.

The recognition of two types of task factors makes it necessary to use different strategies to cope with monotony. Individuals use various strategies to cope with situations. Thus, any intervention or countermeasures might act on any of the marked areas.

This model can be evaluated under certain criteria, similar to the ones applied for theories. One point is that it should not contradict currently held or accepted theories. This aspect has been sufficiently argued in 7.3. It can rather be seen that there is currently no common theory of monotony. The model also contains some kind of taxonomy through the classification of tasks and states. It is easily applicable in ATC through a description of task characteristics and measurable indicators for different states and helps to describe different concepts or working situations. To better evaluate which conditions do predict consistent experimental results, further data needs to be collected. The consideration of positive aspects is something not contained in discussions of monotony and deserves further attention. So, the final scope in a further investigation would be how to balance the factors that a turn of the wheel results in an optimal variation between monotony and positive states.

7.5. CONCLUSIONS UNDER CONSIDERATION OF MONOTONY IN FUTURE ATC CONCEPTS

The general implications of the findings are discussed in the following subsections. The ongoing and planned developments within the Cooperative Air Traffic Management (C-ATM) Operational Concept (Stirnman, Rothmann, Graham, Dowdall, & Eveleigh, 2005) are accompanied with some doubts, as a one-sided focus on complexity and stress issues in ATC leads to develop concepts without sufficiently considering the opposite side of the coin. For this reason, because of the possible negative consequences that appeared in the simulation set-up and the field, the outcomes suggest to further consider these issues in the development of future concepts. In research, work situations dominated by uneventful or repetitive characteristics have resulted in impaired performance and well-being. Also in consideration of incidents that have been reported to occur to a big extent in situations of low or moderate workload the direction of the ongoing development in ATC is worrying. Proposed concepts to deal with traffic increases are not only combined with reduced task variety and control, but impose higher constraints in potential solutions.

With the argument for an increased efficiency in dealing with predicted traffic growth, concepts are currently developed that contain critical factors. This is for example illustrated with the concept of traffic synchronization. Some of the issues, for example the ongoing implementation of data link features, contribute to critical issues as already discussed when investigating the side-effects of automation. Certainly, the concepts are needed to reach the objectives of efficiently dealing with traffic increases; however, some basic human factors issues are to be considered in their development to reduce the potential of evoking monotony. This is similar as for example in the work of Schaefer, Flynn, and Skraaning (2002), where the preferred conflict resolution aid retained the involvement of the controller.

The findings in the study do have implications related to several aspects. First of all, the results affect the design issues. To find out if any of the future concepts has the potential to evoke monotony, concepts should be scanned for the determined factors to be considered in an early step of the task design process. It is emphasized to “keep the controller controlling”, similar to the frequently cited “keep the human-in-the-loop”. This means that when the controller remains the decider for the implemented actions, he or she has also the potential to use the challenge within the traffic through the creation of sufficient task variety. Although, not in all cases decision latitude is an appropriate way to protect operators from developing strain, as a higher level of demands has been reported for such jobs (Rau & Richter, 1996, 279). Richter found a relationship with a high risk of cardiovascular disease in 109 patients whose work was characterized by higher decision latitude compared with a balanced control group. Therefore, careful balance is needed to avoid overload.

Next, the results have some conceptual implications. The oversized centralization of complexity and the implicit solution of reducing complexity is questionable. As Hilburn (2005) mentioned, complexity is related to the difficulty of the traffic. It shall again be emphasized that complexity and activity are not the same, as Zapf (1999) already pointed out. Even in situations of low traffic, a certain task complexity is available in terms of sub-goals, and the action cycle includes a variety of steps to complete the task goals. This is a relevant argument to assess monotony independent of sector complexity.

Finally, the results have implications for the assessment of future concepts. It is not acceptable to conclude only from the assessment of subjective workload if a concept is operationally acceptable. Additional impacts on the subjective and physiological level need to be considered. Nonetheless it is surprising that a state of monotony can result as a probable consequence of repetitive traffic conditions in ATC, especially since up to date research focused predominantly on situations of stress and vigilance.

7.5.1. Recommendations for Dealing with Monotony in ATC Concept Development

The following recommendations for further actions are specifically based on the discussed principles and apply to the research and concept development in ATC. Monotony needs to be considered in the development because smooth transitions from low workload to monotony are likely.

Concerning the conduction of experimental set-ups, the following principles do help to support a decision if a concept contains the potential to evoke monotony:

- Multiple assessments are essential as it reveals opposing developments. This is illustrated in the following example. Results indicated that workload was lower in repetitive than in non repetitive conditions. At the same time, also concentration and attentiveness was lower and monotony was higher. Thus, it is very dangerous to assume that a reduction of workload is a main characteristic for acceptable concepts. Still, currently many projects in the aviation domain are launched with the scope that new procedures or tools should reduce workload. This also indicates the relevance of careful interpretation of contradicting data.

- The collection of very detailed information is necessary to carefully interpret physiological indicators. Important additional information to interpret the data was obtained in interviews and debriefings conducted after the work in the monitored sectors.

- The analysis of individual cases underlines statistically (not) confirmed differences in various work conditions and helps to interpret unexpected results.

- The length of traffic scenarios in a close-to-reality set-up with well-trained controllers on a new concept requires scenario duration of one hour to confirm the effect of monotony. Already after 15 minutes impairments were found on subjective scales, it needs to be considered however that it takes a controller up to 10 minutes to build up the picture of the sector.

- Integration of task analysis in developments of future concepts requires a focus on potential repetitive activities, which is for example addressed in a method proposed by Udris and Alioth (1980).

- Simulations that are addressed to specifically investigate concepts concerning underload aspect need to consider that:

  - with continuous time on task the interest in the simulation might decrease,
  - changes in the environment, e. g. from colleagues can hardly be simulated,
  - risk seeking and search for variety in the activities occur despite clearly defined instructions.

- Finally, every concept is connected with certain potential risks. They need to be defined at an early state of its development to allow the definition of strategies to be applied.

7.6. CONCLUSIONS FOR DEALING WITH MONOTONY IN THE OPERATIONAL ENVIRONMENT

The scope of counteracting monotony is not only to avoid performance problems which might result in measurable incidents, but also to provide well-being and job satisfaction in the working environment, which again has a retro-active loop to the activities during work. The impact of the job satisfaction on the work-related critical states was already indicated in Stokes and Kite (1994, p. 310), who argued that controllers found most stressful the quality of management and administration. This was confirmed by the satisfaction scores obtained in the operational environment and hypothesized relationships with observed increases in satiation that can cover any salient task characteristic that is more likely to evoke monotony because controllers rather discuss the work organization than the task setting.

In relation to the operational environment, two approaches may address work organization on one hand and the individual air traffic controller on the other hand. The strategies can be described for different temporal phases.

The individual factors (e.g., boredom proneness) may be firstly considered in the personnel selection and secondly in the optimization of the controller-task fit. Finally, the ATCO is responsible for his or her functional state, which can be optimized through countermeasures that were included in the experiment. However, there are some open questions related to these issues.

Concerning the strategies it is more likely to fit the individual to the task rather than adapting the task to the individual (as e.g., airspace redesign). Strategies may address the assignment of controllers to working positions, but require the development of more sophisticated tools for supervisors and Flow Management Position (FMP).

The discussion of predictability was already launched, as it contains some risks concerning expectations. Controllers in the centers are often well aware of the risks of expecting low traffic load, but habits might nonetheless play a role in the creation of routine, as this eases the workload a controller has to deal with. Especially in low traffic, controllers often do have the feeling there is nothing to do. This might require consideration in a focused training on the importance of the mental set, as even very low traffic situations do contain some necessary activities (“you always have to check, there is always something that can change”), to be included in an awareness training.

7.6.1. Recommendations for Dealing with Monotony in an Operational Environment

In the following, more specific recommendations are formulated for the operational environments that are separated for the individual and the organization. They are not complete, but do focus on suggestions that can be implemented within a short-term period. The proposed countermeasures also describe agents to implement the countermeasures.

Systematic work position assignment depending on individual initial states (e.g., physical and mental preconditions) and predicted changes in traffic load.

This approach requires improved tools for shift planning applicable by supervisors that allow a more precise prediction of traffic load and adaptation to prior assessed operator states. Especially a systematic assignment to predicted traffic shifts from high to low shall be avoided.

Sector management should not only be based on the traffic count but also consider easily collectable information as the expected departures and arrivals to near airports especially in lower sectors. This helps to avoid underloading conditions if traffic characteristics are perceived of low complexity.

Shifted position assignment.

A shifted exchange with the relief controller for the planning and executive position is a possibility not only to increase the context-related variation, but also contributes to safety as there is always one controller informed about the actual situation in the sector. However, traffic load and subjectively perceived competences of the team colleague need to be considered for an ideal exchange.

Training for ATCOs.

Standardized strategies may be developed and included in training initiatives that address monotony. The role of the mental set (predictability and habits) and controller's awareness of negative side effects of currently used strategies as well as sensitization towards and handling of warning signals announcing critical states in the team colleague are issues that shall be included. Especially simulations may be used to implemented scenarios that increase the awareness of controllers towards their own states and related reactions that go beyond the generally conducted emergency trainings.

Balanced active rest breaks for controllers.

Studies were rare that systematically investigated break and recovery strategies in the ATC environment, even though the relevance of activity is well known by controllers.

Vogt and Leonhardt (2005) reported a positive effect of relaxation techniques, there is however no recommendation concerning physical activity in rest breaks. Under consideration that physical exercises should not expose too high workload, the execution of physical activity is supported and may be transferred even to the controller working at a position in conditions of low traffic load.

Improve incident reporting systems.

A more systematic and detailed collection of information on the controller’s state underlying incidents helps to identify covered causes summarized in commonly used terms such as fatigue and over-confidence especially in low to moderate traffic load.

7.7. OUTLOOK ON FUTURE RESEARCH ON MONOTONY

In the following, several subjects are discussed for further research. It should be kept in mind, that further studies in the field will be necessary not only to replicate the results (Lindsay & Ehrenberg, 1993) but also to better understand the origin of these phenomena. Additional traffic factors may be defined, such as the effect of equal aircraft entry intervals. One of the basic issues might be centered on how to integrate the lessons learnt in systematic approaches not only related to the ATCO, but also to the ANSPs and concept development. This also includes working out a systematic approach to consider effects on critical states implicitly contained in future concepts from the beginning of the development, as the training of mitigation strategies requires specific settings.

Additional research might address if there were specific situations that would be more difficult to solve in monotony, as related to traffic-related loss of information, differences in auditive or visual information presentation or technical problems. Generally, mental set effects or habit interferences have not been systematically addressed. This discussion leads into the role of the cognitive processes for creating an expectation, and might on one hand address the expectation by habits (e.g., the traffic is always like this at this time) and expectation by information (e.g., perception influenced by the principles of gestalt). Both are also related to different temporal courses. The implicit role of monotony in the monitoring activity asks for a more detailed cognitive analysis of the mechanisms how mistakes might occur in repetitive situations. It might be possible that the scanning is becoming less effective in repetitive traffic situations through the expectations how the traffic is.

Finally, the impact of experience is unclear, as two-fold perspectives need to be considered. On one hand, monotony is more likely with increasing experience; on the other hand ATCOs also developed more strategies. From that perspective, monotony might be rather critical for safety after completing the training in the first period as a fully licensed controller. This also relates to the indicated feeling of performing better in higher traffic load. In that respect, an individual factor not considered yet might be the “need for cognition” (Cacioppo, Petty, & Kao, 1984). Defined as the tendency to engage in and enjoy cognitive activities, it might relate to the concept of flow.

Especially with further automation tendencies, this indicator might help in personnel selection and can be considered in combination with the mentioned individual factors in a study using a bigger sample. This does however distract from the primary scope of this research that e is not to replace periods of monotony with conditions evoking stress. It is rather to find an optimal balance between different states that occur and be aware of their negative effects. Still, how much monotony is needed and when is routine imposing a danger for monotony?

Also, currently no statements concerning long-term effects are possible, which would require repeated measurements for a longer period of time. For example, Demerouti et al. (2002) investigated the relationship between short-term strain and long-term burnout in the hospital, and found that those who perceived higher monotony also perceived higher burnout.

The problem of a unified use of terminology was approached with the ISO 10075 for mental workload, which has however not contributed yet to decrease the heterogeneity in expressions that had been in use for a long time.

Overall, a lot of critical aspects have been related to the subject of monotony. Still not much spread is the consideration of positive operator states, expressed in feelings and mood as a resource for ATCOs in efficiently handling traffic. For example, if carefully balanced, monotony may represent a source of recovery or reflection by some ATCOs. This was already recognized by Albert Einstein, who found that “the monotony and solitude of a quiet life stimulates the creative mind”.

Page intentionally left blank

APPENDICES

Page intentionnally left blank

Appendix A: STUDY I

A.1. METHOD

A.1.1. COMPUTATION OF DYNAMIC DENSITY

An adapted indicator of Dynamic Density (DD) was used based on the study of Laudeman et al. (1998) to assess dynamic density in 3-minute-intervals. Because of its relevance in the scenarios the weighting of altitude changes were doubled, the number of AC (AC) was weighted with 0.79 and the number of altitude changes (AL) was weighted (W) with 1.76. As the number of conflict was kept constant in each interval, it was not included in the applied formula, which was simplified to DD=AC x W + AL x W (all elements z-standardized). Table A-1 presents the detailed information for the concerned components of the formula. The summary of the indicators for the various conditions is presented in Table A-2.

Table A-1 Traffic information for scenarios in varied conditions for each interval

a) Scenario repetitive/low dynamic density							
Interval No.	Interval Duration	Conflict Time	Conflict AC	No AC	No Altitude changes	No of conflict	DD (z-score)
1	17:01:50-17:04:50	17:04:50	AFR2444-BAW535	10	2	1	-2.21
2	17:04:50-17:07:50	17:07:50	AFR1045-BAW545	10	3	1	-1.58
3	17:07:50-17:10:50	17:10:50	DLH5648-DAH1419	13	4	1	-.08
4	17:10:50-17:13:50	17:13:50	MLD886-KLM8713	11	4	1	-.65
5	17:13:50-17:16:50	17:16:50	TCX390L-DLH5413	11	4	1	-.65
6	17:16:50-17:19:50	17:19:50	DLH4315-TAP603	11	4	1	-.65
7	17:19:50-17:22:50	17:22:50	CTN508-JKK504	11	4	1	-.65
8	17:22:50-17:25:50	17:25:50	DLH535-AWD565	11	4	1	-.65
9	17:25:50-17:28:50	17:28:50	IBE3247-BAW446	10	4	1	-.94
10	17:28:50-17:31:50	17:31:50	DLH4256-EWG371	12	4	1	-.37
11	17:31:50-17:34:50	17:34:50	AFR5714-BAW565	11	4	1	-.65
12	17:34:50-17:37:50	17:37:50	GOE339-IBE4215	12	4	1	-.37
13	17:37:50-17:40:50	17:40:50	AFR3429-AIH681	14	4	1	.20
14	17:40:50-17:43:50	17:43:50	GAF313-DLH5851	13	4	1	-.08
15	17:43:50-17:46:50	17:46:50	DLH4180-AFR2658	14	4	1	.20
16	17:46:50-17:49:50	17:49:50	DLH5851-RAM851	11	4	1	-.65
b) Scenario repetitive/high dynamic density							
Interval No.	Interval Duration	Conflict Time	Conflict AC	No AC	No Altitude changes	No of conflict	DD (z-score)
1	17:01:50-17:04:50	17:04:50	AZH571-BAW535	10	6	1	.33
2	17:04:50-17:07:50	17:07:50	AFR1046-DLH5413	12	4	1	-.37
3	17:07:50-17:10:50	17:10:50	TRA242-DAH1418	12	5	1	.27
4	17:10:50-17:13:50	17:13:50	SWR545-KLM8713	13	4	1	-.08
5	17:13:50-17:16:50	17:16:50	AIH681-BAW545	12	6	1	.90
6	17:16:50-17:19:50	17:19:50	DLH4315-TAP603	12	7	1	1.54
7	17:19:50-17:22:50	17:22:50	BAW586-JKK604	12	5	1	.27
8	17:22:50-17:25:50	17:25:50	SBE3019-AWD564	12	5	1	.27
9	17:25:50-17:28:50	17:28:50	OAL211-BAW565	12	4	1	-.37
10	17:28:50-17:31:50	17:31:50	AFR5712-EWG371	12	5	1	.27
11	17:31:50-17:34:50	17:34:50	AFR5832-BAW446	12	5	1	.27
12	17:34:50-17:37:50	17:37:50	BVR303-IBE4215	15	5	1	1.13
13	17:37:50-17:40:50	17:40:50	CCH934-AIH673	15	6	1	1.76
14	17:40:50-17:43:50	17:43:50	BRV866-DLH5851	13	5	1	.55
15	17:43:50-17:46:50	17:46:50	MLD884-AFR3648	12	5	1	.27
16	17:46:50-17:49:50	17:49:50	RAM851-GWTIF	12	4	1	-.37

c) Scenario non repetitive/low dynamic density							
Interval No.	Interval Duration	Conflict Time	Conflict AC	No AC	No Altitude changes	No of conflict	DD (z-score)
1	17:01:50-17:04:50	17:04:50	RAM897-SWR332	10	1	1	-2.85
2	17:04:50-17:07:50	17:07:50	KLM1517-RQX854	11	2	1	-1.93
3	17:07:50-17:10:50	17:10:50	IBE4725-RAM323	13	3	1	-.72
4	17:10:50-17:13:50	17:13:50	BAW545-TRA243	14	4	1	.20
5	17:13:50-17:16:50	17:16:50	BMA8715-DAH1419	10	4	1	-.94
6	17:16:50-17:19:50	17:19:50	GOE311-HLF862	11	4	1	-.65
7	17:19:50-17:22:50	17:22:50	DLH5413-AFR573E	12	4	1	-.37
8	17:22:50-17:25:50	17:25:50	TAP603-SWR545	13	4	1	-.08
9	17:25:50-17:28:50	17:28:50	IBE5135-CYP327	13	4	1	-.08
10	17:28:50-17:31:50	17:31:50	CYP327-VPBIE	14	4	1	.20
11	17:31:50-17:34:50	17:34:50	AIH671-TRA245	13	4	1	-.08
12	17:34:50-17:37:50	17:37:50	AZA367-TRA245	14	4	1	.20
13	17:37:50-17:40:50	17:40:50	KLM8713-VKG672	14	4	1	.20
14	17:40:50-17:43:50	17:43:50	AIH681-AWD464	14	4	1	.20
15	17:43:50-17:46:50	17:46:50	KLM1729-RAM851	13	4	1	-.08
16	17:46:50-17:49:50	17:49:50	GMANC-BRT691	10	4	1	-.94

d) Scenario non repetitive/high dynamic density							
Interval No.	Interval Duration	Conflict Time	Conflict AC	No AC	No Altitude changes	No of conflict	DD (z-score)
1	17:01:50-17:04:50	17:04:50	IBE5843-AZA571	9	4	1	-1.22
2	17:04:50-17:07:50	17:07:50	RQX854-IBE4725	13	4	1	-.08
3	17:07:50-17:10:50	17:10:50	DAT3719-BAW545	13	5	1	.55
4	17:10:50-17:13:50	17:13:50	BER7586-AFR452E	14	6	1	1.48
5	17:13:50-17:16:50	17:16:50	AF962ZD-AFR3819	13	7	1	1.83
6	17:16:50-17:19:50	17:19:50	IBE4215-IBE5641	14	5	1	.84
7	17:19:50-17:22:50	17:22:50	IMILA-LOT331	12	5	1	.27
8	17:22:50-17:25:50	17:25:50	BAW446-SBE3019	13	7	1	1.83
9	17:25:50-17:28:50	17:28:50	FIN311-KLM1517	14	6	1	1.48
10	17:28:50-17:31:50	17:31:50	VPBIE-RAM897	15	7	1	2.40
11	17:31:50-17:34:50	17:34:50	AFR3429-JKK504	14	8	1	2.75
12	17:34:50-17:37:50	17:37:50	VEX518-AIH681	13	4	1	-.08
13	17:37:50-17:40:50	17:40:50	KLM1729-GOE313	14	4	1	.20
14	17:40:50-17:43:50	17:43:50	IBE3421-RAM323	12	4	1	-.37
15	17:43:50-17:46:50	17:46:50	MAH550-AFR3429	12	4	1	-.37
16	17:46:50-17:49:50	17:49:50	GMANC-AZA367	11	4	1	-.65

Table A-2 Summary of z-scores for indicators of DD for each scenario condition

DD	Repetitiveness	M	SD	N
low	repetitive	-.61	.61	16
	non repetitive	-.48	.86	16
	Total	-.55	.74	32
high	repetitive	.42	.64	16
	non repetitive	.68	1.17	16
	Total	.55	.93	32
Total	repetitive	-.10	.81	32
	non repetitive	.10	1.17	32
	Total	.00	1.00	64

Note. REP=Repetitiveness; DD=dynamic density[1].

[1]Abbreviations will be used as in Chapter 4 and 5 throughout the appendix, deviations are marked in the note. Deviations from the APA publication guidelines occur.

A.1.2. Description of Different Acknowledge and Vitagraph Analysis Effects

Table A-3 Additional results for data processed with Vitagraph exemplified at HR indicates favorable results for Vitagraph data (compare further Chapter 4.3.1 and Table A-4).

Repetitiveness	RUN	INT	l-h	h-l	l-l	h-h	Total	
			M	M	M	M	M	SD
non repetitive	Run 1	1	8.50	5.95	9.97	7.47	7.97	1.70
		2	10.67	7.07	6.22	6.71	7.67	2.03
		3	10.92	8.04	10.61	6.72	9.07	2.03
		4	9.60	6.69	9.79	7.89	8.49	1.47
		5	6.57	6.27	7.44	7.84	7.03	.73
		6	7.88	5.16	11.49	8.03	8.14	2.60
		7	6.05	4.71	7.62	7.39	6.44	1.35
		8	6.38	2.88	9.64	8.72	6.90	3.01
		9	7.22	4.02	8.22	8.79	7.06	2.13
		10	7.99	4.50	9.68	7.20	7.34	2.16
		11	6.42	4.64	9.22	6.82	6.77	1.88
		12	7.19	3.51	9.22	5.41	6.33	2.44
		13	5.27	4.13	11.44	6.41	6.81	3.22
		14	3.79	2.49	14.04	4.98	6.33	5.24
		15	3.13	2.60	13.21	5.59	6.13	4.90
		16	3.01	2.83	.	4.52	3.45	.93
	Run 2	1	-2.76	1.76	6.32	-.24	1.27	3.84
		2	-.02	1.98	8.01	-.47	2.38	3.90
		3	4.45	1.78	5.44	-.92	2.69	2.86
		4	6.04	2.22	5.94	-.74	3.36	3.26
		5	4.58	3.07	5.31	-1.09	2.97	2.86
		6	2.73	2.20	5.20	-.82	2.33	2.47
		7	1.64	2.07	3.43	-.08	1.76	1.45
		8	4.96	1.98	4.88	1.07	3.22	1.99
		9	4.08	5.93	5.84	-.43	3.86	2.98
		10	2.72	3.86	6.43	1.42	3.61	2.13
		11	4.50	2.13	6.32	1.78	3.68	2.13
		12	1.97	1.33	7.08	2.59	3.24	2.61
		13	2.43	1.82	5.81	2.03	3.02	1.88
		14	4.37	2.06	5.68	.59	3.18	2.28
		15	-.26	1.55	5.72	3.06	2.52	2.53
		16	-1.59	-.66	5.64	1.47	1.22	3.22
repetitive	Run 1	1	5.68	8.39	3.54	3.58	5.30	2.29
		2	6.36	7.96	3.60	3.16	5.27	2.29
		3	7.10	8.52	4.68	5.29	6.40	1.75
		4	5.88	7.79	6.30	5.55	6.38	.99
		5	4.96	7.23	3.40	5.68	5.32	1.59
		6	3.18	6.13	4.14	2.14	3.90	1.70
		7	1.65	6.36	4.03	.43	3.12	2.63
		8	1.78	6.60	2.61	3.24	3.56	2.11
		9	2.81	6.16	3.77	2.12	3.71	1.76
		10	1.44	5.88	4.35	5.01	4.17	1.93
		11	1.31	4.94	3.21	4.80	3.57	1.70
		12	1.41	6.22	2.27	3.09	3.25	2.10
		13	.21	4.93	5.08	4.31	3.63	2.30
		14	.22	5.04	7.08	1.21	3.39	3.22
		15	.89	5.65	6.33	2.64	3.88	2.56
		16	.64	5.05	5.23	1.45	3.09	2.39
	Run 2	1	.62	-4.76	1.39	-.25	-.75	2.75
		2	.95	-2.85	2.83	-1.77	-.21	2.58
		3	.42	-4.45	2.35	1.93	.06	3.12
		4	1.56	-4.41	3.70	.31	.29	3.43
		5	.99	-3.90	3.59	.64	.33	3.11

Repetitiveness	RUN	INT	l-h	h-l	l-l	h-h	Total	
			M	M	M	M	M	SD
		6	1.66	-3.78	2.94	-.78	.01	2.96
		7	1.81	-3.66	4.37	-.60	.48	3.42
		8	.42	-4.54	3.49	-.50	-.28	3.31
		9	.21	-3.84	3.36	.09	-.04	2.95
		10	.27	-2.15	3.48	-.14	.37	2.33
		11	1.65	-3.56	4.66	-2.49	.06	3.80
		12	.91	.20	3.74	-1.79	.76	2.29
		13	.20	-2.28	3.16	-1.31	-.06	2.38
		14	.03	-3.02	1.30	-1.67	-.84	1.90
		15	-.22	-3.50	2.42	-1.22	-.63	2.45
		16	.24	-4.11	4.54	-4.26	-.90	4.18

Table A-4: Results of Analysis of Variance for HR analyzed with Vitagraph

SOURCE	RESULTS (F_{df}, p-value)
	HR (baseline corr.) with Vitagraph
REP	$F_{1,2}$=77.22; p=.013*
DD	$F_{3,2}$=13.17; p=.071(*)
RUN	$F_{1,2}$=5.87; p=.136(*)
INTER	$F_{15,30}$=2.35; p=.023*
Run x Rep	$F_{1,2}$=.09; p=.790
Run x DD	$F_{3,2}$=.49; p=.722
Inter x Rep	$F_{15,30}$=.59; p=.859
Inter x DD	$F_{45,30}$=1.01; p=.498
Run x Inter	$F_{15,30}$=2.29; p=.026*
Run x Inter x Rep	$F_{15,30}$=.49; p=.924
Run x Inter x DD	$F_{45,30}$=.84; p=.703

Note. N=8. Rep=Repetitiveness; DD=Sequence of DD; Inter=Interval during run. ***p<.001. **p<.01. *p<.05. (*)p<0.2.

Table A-5: Settings for Vitaport Channels (PP=Preprocessing; Res=Resolution; Amp=amplification;)

Signal	Sample (Hz)	Store (Hz)	Unit	MUL/DIV	Offset	amp	HP in sec	LP
ECG	256	128	mV	2888/30000	2048	1505.4	0.15	149.7
EDA	16	8	µS	-	-	-	D.C.	-
EOG_H	256	256	mV	2888/30000	2048	1505.4	5.000	49.7
EOG_V	256	256	mV	2888/30000	2048	1505.4	5.000	49.7
CZ	256	256	µV	2888/30000	2048	5000	3.000	70
Resp	256	16	adc	1/1	2047	100	5.000	300
Move	256	32	g	2/1325	2048	20	D.C.	149.7
marker			adc					

A.2. DESCRIPTIVE STATISTICS AND ADDITIONAL RESULTS OF STATISTICAL ANALYSIS

A.2.1. Results for Physiological Data

Table A-6: Descriptive Statistics: HR (corr.)

RUN	Run 1	DD	l-h	h-l	l-l	h-h	Total	
REP	Repetitive	Interval	M	M	M	M	M	SD
		1	4.92	10.58	3.49	3.48	5.62	3.38
		2	5.64	9.95	3.51	3.18	5.57	3.12
		3	6.33	10.58	4.58	5.16	6.66	2.71
		4	5.20	9.92	6.23	5.37	6.68	2.21
		5	4.22	9.22	3.31	5.51	5.56	2.60
		6	2.42	8.34	4.06	2.07	4.22	2.88
		7	.99	8.31	3.98	.73	3.50	3.53
		8	1.10	8.87	2.57	3.14	3.92	3.41
		9	2.13	8.28	3.71	2.12	4.06	2.91
		10	.77	8.16	4.25	4.78	4.49	3.03
		11	.59	6.99	3.18	4.68	3.86	2.69
		12	.73	8.24	2.24	2.96	3.54	3.27
		13	-.50	7.01	4.94	4.19	3.91	3.17
		14	-.49	7.01	6.93	1.21	3.67	3.88
		15	.27	7.92	6.17	2.55	4.23	3.46
		16	-.08	7.21	5.06	1.45	3.41	3.32

RUN	Run 1	DD	l-h	h-l	l-l	h-h	Total	
REP	Non Repet.	Interval	M	M	M	M	M	SD
		1	8.09	5.82	9.78	7.43	7.78	1.64
		2	10.16	6.94	6.13	6.62	7.46	1.83
		3	10.41	7.90	10.42	6.66	8.85	1.88
		4	9.09	6.57	9.60	7.84	8.28	1.36
		5	6.25	6.14	7.37	7.65	6.85	.77
		6	7.49	5.02	11.48	7.93	7.98	2.66
		7	5.65	4.64	7.51	7.38	6.30	1.39
		8	5.96	2.83	9.48	8.61	6.72	2.99
		9	6.88	3.95	8.09	8.79	6.93	2.14
		10	7.43	4.39	9.93	6.73	7.12	2.28
		11	6.20	4.56	9.08	6.84	6.67	1.87
		12	6.75	3.42	9.11	5.19	6.12	2.41
		13	4.87	4.04	11.20	6.64	6.69	3.20
		14	3.45	2.46	13.73	4.50	6.04	5.20
		15	2.83	2.61	12.95	5.57	5.99	4.83
		16	2.65	2.75	12.95	4.44	5.70	4.90

RUN	Run 2	DD	l-h	h-l	l-l	h-h	Total	
REP	Repetitive	Interval	M	M	M	M	M	SD
		1	-.06	-2.11	1.36	-.02	-.21	1.43
		2	.25	-.57	2.82	.36	.71	1.46
		3	-.28	-1.94	2.34	1.95	.52	2.01
		4	.83	-2.06	3.61	.41	.70	2.32
		5	.35	-1.49	3.57	.95	.84	2.09
		6	.95	-1.46	2.96	-.09	.59	1.86
		7	1.13	-1.25	4.27	.02	1.04	2.36

RUN	Run 2	DD	l-h	h-l	l-l	h-h	Total	
REP	Repetitive	Interval	M	M	M	M	M	SD
		8	-.23	-2.21	3.45	.00	.25	2.35
		9	-.47	-1.50	3.27	.06	.34	2.06
		10	-.37	.10	3.42	-.04	.78	1.77
		11	.97	-1.29	4.68	-.17	1.05	2.59
		12	.18	2.46	3.65	-1.43	1.22	2.27
		13	-.46	-.06	3.01	-.87	.41	1.77
		14	-.62	-.71	1.25	-1.08	-.29	1.05
		15	-.83	-1.18	2.40	-1.12	-.19	1.73
		16	-.50	-1.85	4.48	.63	.69	2.72

RUN	Run 2	DD	l-h	h-l	l-l	h-h	Total	
REP	Non Repet.	Interval	M	M	M	M	M	SD
		1	-2.99	1.75	6.33	-.22	1.22	3.93
		2	-.33	1.97	7.88	-.34	2.29	3.88
		3	4.57	1.77	5.42	-.87	2.72	2.85
		4	5.90	2.18	5.90	-.64	3.33	3.18
		5	4.26	2.94	5.28	-1.02	2.86	2.76
		6	2.49	2.38	5.10	-.75	2.31	2.40
		7	1.40	2.07	3.58	-.09	1.74	1.52
		8	4.52	2.21	4.81	1.13	3.17	1.79
		9	3.82	5.82	5.75	-.42	3.74	2.92
		10	2.40	3.77	6.35	1.43	3.49	2.14
		11	4.18	2.10	6.25	1.80	3.58	2.07
		12	1.63	1.53	6.99	2.55	3.17	2.58
		13	2.21	1.75	5.76	2.00	2.93	1.89
		14	4.25	2.07	5.69	.66	3.17	2.24
		15	-.56	1.52	5.64	3.01	2.40	2.61
		16	-1.84	-.75	5.55	1.47	1.10	3.27

Table A-7: Descriptive Statistics: HRV

RUN	Run 1	DD	l-h	h-l	l-l	h-h	Total	
REP	Repetitive	Interval	M	M	M	M	M	SD
		1	5.11	17.81	30.16	7.95	15.26	11.33
		2	7.40	11.40	27.17	11.28	14.31	8.77
		3	7.22	10.88	25.25	8.79	13.03	8.28
		4	7.83	18.94	39.49	10.28	19.14	14.38
		5	6.10	12.52	34.44	13.80	16.72	12.29
		6	5.68	15.67	31.51	15.55	17.10	10.69
		7	4.73	17.75	54.38	5.29	20.54	23.35
		8	6.86	13.03	22.69	12.78	13.84	6.55
		9	4.77	11.36	21.05	6.35	10.88	7.34
		10	6.02	26.57	11.95	14.11	14.66	8.64
		11	8.73	18.90	27.60	9.33	16.14	8.95
		12	5.10	14.66	25.80	9.93	13.87	8.86
		13	5.59	15.54	29.00	10.54	15.17	10.08
		14	4.30	12.70	33.41	9.25	14.91	12.81
		15	5.45	11.43	34.92	18.09	17.47	12.73
		16	2.93	13.07	28.78	8.62	13.35	11.09

RUN	Run 1		l-h	h-l	l-l	h-h	Total	
REP	Non repet.	Interval	M	M	M	M	M	SD
		1	6.59	4.81	25.17	9.18	11.44	9.33
		2	5.63	6.27	20.09	9.06	10.26	6.72
		3	15.31	6.61	24.44	13.55	14.98	7.34
		4	14.08	5.46	21.56	10.51	12.90	6.77
		5	14.61	4.28	22.63	13.83	13.84	7.51
		6	9.30	6.24	40.33	11.55	16.85	15.80

		7	8.19	10.19	31.34	11.30	15.26	10.80
		8	8.33	5.06	37.19	17.45	17.01	14.44
		9	13.02	9.24	36.09	17.63	18.99	11.90
		10	12.64	6.76	53.86	13.96	21.81	21.60
		11	15.98	9.64	35.72	11.48	18.21	11.98
		12	10.79	9.28	25.38	9.65	13.78	7.76
		13	9.60	6.49	16.37	12.88	11.33	4.25
		14	7.21	7.52	14.33	12.19	10.31	3.52
		15	12.96	6.19	22.78	28.83	17.69	10.08
		16	10.61	9.13	22.78	21.34	15.96	7.09

RUN	Run 2		l-h	h-l	l-l	h-h	Total	
REP	Repetitive	Interval	M	M	M	M	M	SD
		1	5.03	12.15	23.41	20.60	15.30	8.35
		2	5.03	12.15	23.41	20.60	15.30	8.35
		3	4.31	14.78	17.56	12.53	12.30	5.71
		4	4.80	8.92	18.75	8.01	10.12	6.02
		5	5.07	12.36	30.80	11.65	14.97	11.05
		6	3.83	18.59	20.22	10.87	13.38	7.56
		7	7.60	21.61	23.21	9.33	15.44	8.11
		8	6.89	16.32	35.30	9.90	17.10	12.76
		9	4.74	5.61	25.46	4.05	9.96	10.35
		10	6.22	19.13	26.34	8.38	15.02	9.43
		11	5.74	14.30	30.01	3.39	13.36	12.05
		12	6.97	18.00	35.51	8.76	17.31	13.06
		13	4.51	14.20	38.36	14.41	17.87	14.42
		14	7.00	12.52	22.65	8.04	12.55	7.14
		15	5.33	12.55	27.13	10.53	13.89	9.34
		16	8.81	16.46	38.15	10.33	18.44	13.55

RUN	Run 2		l-h	h-l	l-l	h-h	Total	
REP	Non repet.	Interval	M	M	M	M	M	SD
		1	8.69	8.69	31.72	14.95	16.01	10.88
		2	11.41	5.69	30.08	17.08	16.07	10.43
		3	19.83	6.61	24.42	11.56	15.60	8.02
		4	9.95	9.98	23.61	10.07	13.40	6.81
		5	19.30	6.34	31.97	12.74	17.59	10.95
		6	8.03	10.18	25.53	12.52	14.07	7.86
		7	10.47	6.57	42.43	15.58	18.76	16.20
		8	10.82	8.57	27.79	19.43	16.65	8.78
		9	10.64	9.08	32.20	13.01	16.24	10.77
		10	14.76	7.87	16.44	11.67	12.68	3.77
		11	15.24	9.45	23.07	15.44	15.80	5.59
		12	13.43	9.65	36.61	17.06	19.19	12.00
		13	12.67	9.20	28.27	13.81	15.99	8.42
		14	22.42	9.21	34.98	15.97	20.64	10.97
		15	12.27	6.78	26.93	12.42	14.60	8.63
		16	23.13	7.44	28.77	11.75	17.77	9.88

Table A-8: Descriptive Statistics: SCL (corr.)

RUN	Run 1		l-h	h-l	l-l	h-h	Total	
REP	Repetitive	Interval	M	M	M	M	M	SD
		1	.56	.80	.51	.75	.66	.14
		2	.52	.66	.47	.61	.56	.08
		3	.58	.62	.38	.52	.52	.11
		4	.47	.60	.36	.48	.48	.10
		5	.36	.69	.44	.43	.48	.15
		6	.41	.63	.33	.43	.45	.13
		7	.27	.57	.41	.52	.44	.13
		8	.34	.61	.46	.42	.46	.11
		9	.27	.48	.40	.36	.38	.09
		10	.19	.42	.48	.31	.35	.13
		11	.19	.39	.35	.27	.30	.09
		12	.15	.62	.37	.27	.35	.20
		13	.19	.64	.21	.24	.32	.22
		14	.20	.49	.15	.21	.26	.16
		15	.26	.44	.12	.18	.25	.14
		16	.27	.46	.11	.16	.25	.15

RUN	Run 1		l-h	h-l	l-l	h-h	Total	
REP	Non repet.	Interval	M	M	M	M	M	SD
		1	.56	.80	.51	.75	.66	.14
		2	.52	.66	.47	.61	.56	.08
		3	.58	.62	.38	.52	.52	.11
		4	.47	.60	.36	.48	.48	.10
		5	.36	.69	.44	.43	.48	.15
		6	.41	.63	.33	.43	.45	.13
		7	.27	.57	.41	.52	.44	.13
		8	.34	.61	.46	.42	.46	.11
		9	.27	.48	.40	.36	.38	.09
		10	.19	.42	.48	.31	.35	.13
		11	.19	.39	.35	.27	.30	.09
		12	.15	.62	.37	.27	.35	.20
		13	.19	.64	.21	.24	.32	.22
		14	.20	.49	.15	.21	.26	.16
		15	.26	.44	.12	.18	.25	.14
		16	.27	.46	.11	.16	.25	.15

RUN	Run 2		l-h	h-l	l-l	h-h	Total	
REP	Repetitive	Interval	M	M	M	M	M	SD
		1	.20	.37	.39	.17	.28	.11
		2	.16	.35	.26	.21	.24	.08
		3	.10	.26	.25	.18	.20	.08
		4	.05	.25	.17	.14	.15	.08
		5	.01	.21	.12	.12	.11	.08
		6	.06	.10	.05	.09	.08	.03
		7	.06	.09	.14	.09	.09	.04
		8	.05	.06	.34	.06	.13	.14
		9	.02	.06	.33	.06	.12	.15
		10	-.02	.07	.31	.07	.11	.14
		11	-.03	.05	.42	.05	.12	.20
		12	.02	.07	.53	.07	.17	.24
		13	-.01	.06	.49	.06	.15	.23
		14	-.03	.04	.49	.04	.14	.24
		15	-.04	.02	.33	.02	.08	.17
		16	.00	.01	.25	.01	.07	.12

RUN	Run 2		l-h	h-l	l-l	h-h	Total	

REP	Non repet.	Interval	M	M	M	M	M	SD
		1	.21	.22	.85	.88	.54	.38
		2	.13	.16	.88	.86	.51	.42
		3	.16	.20	.81	.80	.49	.36
		4	.27	.15	.80	.63	.46	.30
		5	.22	.13	.81	.44	.40	.30
		6	.17	.09	.89	.31	.37	.36
		7	.14	.11	.87	.33	.36	.35
		8	.19	.03	.90	.36	.37	.38
		9	.26	.09	.90	.25	.37	.36
		10	.27	.06	.94	.12	.35	.40
		11	.25	.05	.83	.16	.32	.35
		12	.20	.06	.88	.30	.36	.36
		13	.11	.03	.90	.40	.36	.40
		14	.16	-.04	.89	.42	.36	.40
		15	.18	-.02	.86	.46	.37	.39
		16	.16	-.03	.83	.42	.34	.37

Table A-9: Descriptive Statistics: No. of blinks

RUN	Run 1		l-h	h-l	l-l	h-h	Total	
REP	Repetitive	Interval	M	M	M	M	M	SD
		1	4.00	52.00	28.00	25.00	27.25	19.65
		2	9.00	72.00	21.00	19.00	30.25	28.32
		3	6.00	72.00	21.00	21.00	30.00	28.88
		4	15.00	48.00	30.00	16.00	27.25	15.44
		5	3.00	61.00	25.00	24.00	28.25	24.07
		6	7.00	52.00	32.00	20.00	27.75	19.12
		7	10.00	50.00	19.00	20.00	24.75	17.42
		8	11.00	55.00	33.00	20.00	29.75	19.10
		9	10.00	52.00	28.00	22.00	28.00	17.66
		10	11.00	72.00	40.00	34.00	39.25	25.16
		11	24.00	66.00	12.00	27.00	32.25	23.41
		12	9.00	54.00	12.00	18.00	23.25	20.84
		13	14.00	62.00	20.00	26.00	30.50	21.56
		14	9.00	64.00	30.00	24.00	31.75	23.24
		15	18.00	58.00	18.00	26.00	30.00	19.04
		16	4.00	52.00	28.00	25.00	27.25	19.65

RUN	Run 2		l-h	h-l	l-l	h-h	Total	
REP	Non repet.	Interval	M	M	M	M	M	SD
		1	37.00	22.00	63.00	25.00	36.75	18.66
		2	26.00	32.00	67.00	21.00	36.50	20.82
		3	26.00	26.00	62.00	19.00	33.25	19.45
		4	27.00	22.00	58.00	21.00	32.00	17.53
		5	40.00	29.00	64.00	16.00	37.25	20.35
		6	37.00	19.00	74.00	17.00	36.75	26.41
		7	33.00	23.00	69.00	23.00	37.00	21.85
		8	30.00	20.00	60.00	17.00	31.75	19.64
		9	25.00	28.00	63.00	22.00	34.50	19.16
		10	32.00	18.00	59.00	22.00	32.75	18.46
		11	33.00	14.00	62.00	19.00	32.00	21.56
		12	33.00	21.00	78.00	17.00	37.25	28.00
		13	34.00	23.00	61.00	17.00	33.75	19.48
		14	32.00	21.00	54.00	22.00	32.25	15.33
		15	47.00	22.00	74.00	16.00	39.75	26.49
		16	56.00	24.00	74.00	19.00	43.25	26.25

RUN	Run 2		l-h	h-l	l-l	h-h	Total	
REP	repetitive	Interval	M	M	M	M	M	SD
		1	10.00	75.00	20.00	35.00	35.00	28.58
		2	7.00	68.00	24.00	32.00	32.75	25.71
		3	18.00	51.00	36.00	30.00	33.75	13.72
		4	17.00	54.00	23.00	23.00	29.25	16.74
		5	9.00	53.00	23.00	24.00	27.25	18.48
		6	6.00	57.00	31.00	29.00	30.75	20.85
		7	10.00	63.00	24.00	32.00	32.25	22.43
		8	3.00	47.00	24.00	28.00	25.50	18.05
		9	18.00	66.00	24.00	25.00	33.25	22.05
		10	9.00	59.00	27.00	23.00	29.50	21.13
		11	7.00	55.00	23.00	18.00	25.75	20.61
		12	6.00	71.00	16.00	17.00	27.50	29.42
		13	4.00	54.00	28.00	33.00	29.75	20.53
		14	19.00	48.00	34.00	32.00	33.25	11.87
		15	18.00	59.00	34.00	27.00	34.50	17.60
		16	14.00	62.00	26.00	31.00	33.25	20.45

RUN	Run 2		l-h	h-l	l-l	h-h	Total	Total
REP	Non repet.	Interval	M	M	M	M	M	SD
		1	36.00	27.00	64.00	24.00	37.75	18.23
		2	24.00	36.00	62.00	17.00	34.75	19.79
		3	27.00	18.00	66.00	20.00	32.75	22.50
		4	20.00	21.00	62.00	21.00	31.00	20.67
		5	31.00	24.00	58.00	19.00	33.00	17.38
		6	20.00	26.00	74.00	22.00	35.50	25.79
		7	22.00	27.00	79.00	19.00	36.75	28.36
		8	21.00	18.00	68.00	24.00	32.75	23.63
		9	44.00	29.00	63.00	24.00	40.00	17.53
		10	31.00	17.00	65.00	25.00	34.50	21.13
		11	26.00	24.00	77.00	20.00	36.75	26.95
		12	18.00	19.00	64.00	22.00	30.75	22.23
		13	20.00	28.00	65.00	25.00	34.50	20.60
		14	42.00	27.00	54.00	27.00	37.50	13.08
		15	57.00	39.00	73.00	23.00	48.00	21.69
		16	72.00	40.00	69.00	27.00	52.00	22.05

A.2.2. Results for Subjective Data

Table A-10: Descriptive Statistics: Items based on Thackray et al. (1975)

Attentiveness		DD	l-h	h-l	l-l	h-h	Total	
Repetitiveness	RUN	Interval	M	M	M	M	M	SD
repetitive	Run 1	1	.13	1.75	.38	.13	.59	.78
		2	.13	-.25	.38	1.13	.34	.58
		3	.13	-.25	.38	.13	.09	.26
	Run 2	1	.13	.75	.38	.13	.34	.30
		2	.13	-.25	.38	-.88	-.16	.54
		3	.13	-1.25	.38	.13	-.16	.74
Non repetitive	Run 1	1	.25	.88	.75	.53	.33	.25
		2	.25	-1.13	.75	.03	.81	.25
		3	-.75	-.13	-.25	-.22	.41	.25
	Run 2	1	.25	-.13	.75	.28	.36	.25
		2	.25	-.13	-.25	.03	.26	.25
		3	.25	.88	.75	.53	.33	.25

Fatigue		DD	l-h	h-l	l-l	h-h	Total	
Repetitiveness		Interval	M	M	M	M	M	SD
repetitive	Run 1	1	2.00	-.75	-.25	-.38	.16	1.25
		2	.00	-.75	.75	-.38	-.09	.64
		3	.00	.25	-.25	.63	.16	.37
	Run 2	1	.00	.25	.75	-.38	.16	.47
		2	-1.00	.25	.75	.63	.16	.80
		3	-1.00	1.25	.75	.63	.41	.98
Non repetitive	Run 1	1	-1.38	-1.13	-.63	.75	-.59	.95
		2	-.38	-.13	-.63	-.25	-.34	.21
		3	1.63	.88	-.63	.75	.66	.94
	Run 2	1	.63	-1.13	.38	-.25	-.09	.78
		2	-.38	-.13	1.38	-.25	.16	.82
		3	-.38	.88	.38	-.25	.16	.58

Boredom		DD	l-h	h-l	l-l	h-h	Total	
Repetitiveness		Interval	M	M	M	M	M	SD
repetitive	Run 1	1	.50	-.88	-.50	-.75	-.41	.62
		2	.50	-.88	-.50	.25	-.16	.64
		3	.50	.13	.50	-.50	.09	.59
	Run 2	1	-.50	.13	.50	-.50	-.16	.57
		2	-.50	1.13	-.50	.50	.34	.98
		3	-.50	1.13	-.50	.50	.34	.98
Non repetitive	Run 1	1	-.13	-.88	2.75	-.50	.31	1.65
		2	-.13	.13	-.25	-.50	-.19	.26
		3	-.13	.13	-1.25	-.50	-.44	.60
	Run 2	1	-.13	.13	-.25	-.50	-.19	.26
		2	-.13	.13	-.25	2.50	.56	1.30
		3	.88	.13	.75	.50	.56	.33

Irritation		DD	l-h	h-l	l-l	h-h	Total	
Repetitiveness		Interval	M	M	M	M	M	SD
repetitive	Run 1	1	.00	-.50	-.38	.63	.00	.62
		2	.00	-.50	-.38	.63	.00	.47
		3	.00	-.50	-.38	-.38	-.25	.23
	Run 2	1	.00	-.50	-.38	.63	.00	.47
		2	.00	1.50	-.38	.63	.25	.85
		3	.00	.50	-.38	-.38	.00	.37
Non repetitive	Run 1	1	-.38	-1.25	-.38	-1.50	-.88	.59
		2	-.38	.75	.63	.50	.38	.51
		3	-.38	1.75	.63	2.50	1.13	1.26
	Run 2	1	.63	-.25	-.38	.50	.13	.51
		2	1.63	-.25	.63	1.50	.88	.87
		3	.63	-.25	-.38	-.50	-.13	.51

Strain		DD	l-h	h-l	l-l	h-h	Total	
Repetitiveness		Interval	M	M	M	M	M	SD
repetitive	Run 1	1	.50	-.88	1.13	.88	.41	.89
		2	.50	.13	.13	-.13	.16	.26
		3	-.50	-.88	-.88	-.13	-.59	.36
	Run 2	1	.50	.13	1.13	-.13	.41	.54
		2	-.50	1.13	-.88	-.13	-.09	.87
		3	-.50	.13	1.13	-.13	.16	.70
Non repetitive	Run 1	1	1.88	-1.00	.88	.38	.53	1.20
		2	1.88	1.00	-.13	.38	.78	.86
		3	-.13	1.00	.88	.38	.53	.51
	Run 2	1	-.13	1.00	-.13	-.63	.03	.69
		2	-.13	-1.00	-.13	.38	-.22	.57

		3	-1.13	.00	-.13	.38	-.22	.64

Concentration		DD	l-h	h-l	l-l	h-h	Total	
Repetitiveness		Interval	M	M	M	M	M	SD
repetitive	Run 1	1	.00	1.38	.38	.25	.50	.60
		2	.00	.38	.38	1.25	.50	.53
		3	.00	-.63	.38	.25	.00	.44
	Run 2	1	.00	.38	.38	.25	.25	.18
		2	.00	.38	.38	-.75	.00	.53
		3	.00	-.63	.38	.25	.00	.44
Non repetitive	Run 1	1	.25	.13	1.00	-.13	.31	.48
		2	.25	.13	.00	.88	.31	.39
		3	.25	-.88	1.00	-.13	.06	.78
	Run 2	1	.25	.13	1.00	-.13	.31	.48
		2	.25	.13	.00	.88	.31	.39
		3	.25	.13	-2.00	.88	-.19	1.25

Motivation		DD	l-h	h-l	l-l	h-h	Total	
Repetitiveness		Interval	M	M	M	M	M	SD
repetitive	Run 1	1	-.13	1.25	.25	.38	.44	.58
		2	-.13	.25	.25	.38	.19	.22
		3	-.13	-.75	.25	.38	-.06	.51
	Run 2	1	-.13	-.75	.25	.38	-.06	.51
		2	-.13	-.75	.25	-.63	-.31	.46
		3	-.13	-.75	.25	-.63	-.31	.46
Non repetitive	Run 1	1	.25	.00	1.63	.25	.53	.74
		2	.25	.00	-.38	-.75	-.22	.44
		3	.25	.00	.63	.25	.28	.26
	Run 2	1	.25	.00	-.38	.25	.03	.30
		2	.25	.00	-.38	.25	.03	.30
		3	.25	.00	-1.38	.25	-.22	.78

Sleepiness		DD	l-h	h-l	l-l	h-h	Total	
Repetitiveness		Interval	M	M	M	M	M	SD
repetitive	Run 1	1	.13	-1.13	-1.63	-.13	-.69	.83
		2	.13	-1.13	-1.63	-.13	-.69	.83
		3	.13	-.13	.38	-.13	.06	.24
	Run 2	1	-.88	-.13	1.38	-.13	.06	.94
		2	-.88	.88	1.38	-.13	.31	1.01
		3	.13	.88	1.38	.88	.81	.52
Non repetitive	Run 1	1	-1.25	-1.38	.50	.13	-.81	.60
		2	-1.25	-.38	.50	.13	-.31	.73
		3	.75	.63	.50	.13	.19	.60
	Run 2	1	-.25	-1.38	.50	.13	-.31	.78
		2	-.25	-.38	.50	.13	-.06	.39
		3	-.25	.63	.50	.13	.19	.44

Table A-11: Descriptive Statistics: Workload assessment (Nasa-TLX) + feeling of monotony

Repetitiveness	DD	l-h	h-l	l-l	h-h	Total	
repetitive	RUN	M	M	M	M	M	SD
Mental demand	Run 1	42.00	30.00	62.00	57.00	47.75	14.57
	Run 2	23.00	25.00	63.00	28.00	34.75	18.95
Physical demand	Run 1	8.00	10.00	13.00	48.00	19.75	18.95
	Run 2	8.00	15.00	38.00	28.00	22.25	13.38
Temporal demand	Run 1	28.00	10.00	17.00	42.00	24.25	13.96
	Run 2	23.00	13.00	38.00	18.00	23.00	10.80
Performance	Run 1	28.00	30.00	48.00	42.00	37.00	9.59
	Run 2	33.00	20.00	53.00	18.00	31.00	16.10
Effort	Run 1	18.00	25.00	2.00	28.00	18.25	11.62
	Run 2	23.00	45.00	3.00	18.00	22.25	17.39
Frustration	Run 1	8.00	40.00	3.00	28.00	19.75	17.29
	Run 2	18.00	48.00	23.00	18.00	26.75	14.36
Feeling of monotony	Run 1	72.00	60.00	67.00	37.00	59.00	15.47
	Run 2	58.00	87.00	77.00	57.00	69.75	14.73

Repetitiveness	DD	l-h	h-l	l-l	h-h	Total	
non repetitive	RUN	M	M	M	M	M	SD
Mental demand	Run 1	68.00	80.00	52.00	38.00	59.50	18.36
	Run 2	93.00	60.00	63.00	38.00	63.50	22.61
Physical demand	Run 1	13.00	50.00	12.00	28.00	25.75	17.75
	Run 2	23.00	20.00	33.00	40.00	29.00	9.20
Temporal demand	Run 1	63.00	70.00	13.00	28.00	43.50	27.40
	Run 2	93.00	30.00	52.00	30.00	51.25	29.70
Performance	Run 1	63.00	60.00	32.00	29.00	46.00	17.98
	Run 2	78.00	50.00	62.00	29.00	54.75	20.65
Effort	Run 1	18.00	10.00	17.00	70.00	28.75	27.73
	Run 2	78.00	10.00	67.00	20.00	43.75	33.75
Frustration	Run 1	13.00	30.00	38.00	60.00	35.25	19.52
	Run 2	62.00	10.00	82.00	50.00	51.00	30.35
Feeling of monotony	Run 1	13.00	10.00	37.00	40.00	25.00	15.68
	Run 2	13.00	20.00	18.00	40.00	22.75	11.87

Table A-12: Descriptive Statistics: Mood Assessment (UWIST)

Repetitiveness		DD	l-h	h-l	l-l	h-h	Total	
		RUN	M	M	M	M	M	SD
repetitive	Hedonic tone	Run 1	3.25	2.75	3.25	3.13	3.09	.24
		Run 2	3.50	3.25	3.88	3.13	3.44	.33
	Tense arousal	Run 1	1.13	1.50	1.38	2.13	1.53	.43
		Run 2	1.50	1.88	1.25	2.13	1.69	.39
	Energetic arousal	Run 1	2.38	3.00	3.00	2.63	2.75	.31
		Run 2	3.00	2.88	3.00	2.38	2.81	.30
non repetitive	Hedonic tone	Run 1	3.25	3.63	3.25	3.25	3.34	.19
		Run 2	2.50	3.75	2.38	3.38	3.00	.67
	Tense arousal	Run 1	2.00	1.88	2.25	2.25	2.09	.19
		Run 2	2.13	1.50	2.00	2.25	1.97	.33
	Energetic arousal	Run 1	2.50	2.75	2.75	3.13	2.78	.26
		Run 2	2.25	2.50	2.25	2.75	2.44	.24

Table A-13: Descriptive Statistics: Strain Assessment (SOF)

Repetitiveness		DD RUN	l-h M	h-l M	l-l M	h-h M	Total M	Total SD
repetitive	Stress	Run 1	1.40	1.50	1.50	2.10	1.63	.32
		Run 2	1.50	1.50	1.60	1.89	1.62	.18
	Satiation	Run 1	1.44	1.56	1.67	1.78	1.61	.14
		Run 2	1.56	2.22	1.89	1.56	1.81	.32
	Fatigue	Run 1	1.50	1.80	1.70	2.10	1.78	.25
		Run 2	1.60	2.40	1.80	1.70	1.88	.36
	Monotony	Run 1	2.00	2.10	1.80	2.20	2.03	.17
		Run 2	2.20	2.60	2.00	2.40	2.30	.26
non repetitive	Stress	Run 1	2.50	1.90	2.00	1.80	2.05	.31
		Run 2	2.20	1.90	2.60	1.80	2.13	.36
	Satiation	Run 1	2.22	1.00	1.22	1.33	1.44	.54
		Run 2	2.33	1.67	2.33	1.56	1.97	.42
	Fatigue	Run 1	2.00	1.60	1.60	1.90	1.78	.21
		Run 2	2.10	1.80	2.44	1.90	2.06	.28
	Monotony	Run 1	2.50	1.90	1.60	1.60	1.90	.42
		Run 2	2.30	1.90	2.00	1.60	1.95	.29

Table A-14: Descriptive Statistics: Situation Awareness (SASHA)

Repetitiveness	Item	DD RUN	l-h M	h-l M	l-l M	h-h M	Total M	Total SD
repetitive	1	Run 1	5.00	4.00	5.00	5.00	4.75	.50
		Run 2	5.00	4.00	5.00	5.00	4.75	.50
	2	Run 1	5.00	4.00	5.00	5.00	4.75	.50
		Run 2	5.00	4.00	5.00	5.00	4.75	.50
	3	Run 1	1.00	2.00	1.00	2.00	1.50	.58
		Run 2	3.00	2.00	2.00	1.00	2.00	.82
	4	Run 1	1.00	1.00	1.00	1.00	1.00	.00
		Run 2	1.00	1.00	1.00	2.00	1.25	.50
	5	Run 1	2.00	1.00	1.00	1.00	1.25	.50
		Run 2	1.00	1.00	1.00	1.00	1.00	.00
	6	Run 1	1.00	1.00	2.00	1.00	1.25	.50
		Run 2	1.00	1.00	1.00	1.00	1.00	.00
	7	Run 1	1.00	2.00	1.00	1.00	1.25	.50
		Run 2	1.00	1.00	1.00	1.00	1.00	.00
	8	Run 1	5.00	4.00	5.00	4.00	4.50	.58
		Run 2	5.00	3.00	5.00	4.00	4.25	.96
non repetitive	1	Run 1	4.00	4.00	4.00	4.00	4.00	.00
		Run 2	4.00	5.00	3.00	5.00	4.25	.96
	2	Run 1	4.00	5.00	4.00	4.00	4.25	.50
		Run 2	4.00	5.00	4.00	5.00	4.50	.58
	3	Run 1	2.00	3.00	2.00	1.00	2.00	.82
		Run 2	1.00	1.00	4.00	1.00	1.75	1.50
	4	Run 1	1.00	1.00	1.00	2.00	1.25	.50
		Run 2	1.00	1.00	4.00	1.00	1.75	1.50
	5	Run 1	3.00	1.00	2.00	1.00	1.75	.96
		Run 2	3.00	1.00	2.00	1.00	1.75	.96
	6	Run 1	3.00	1.00	2.00	1.00	1.75	.96
		Run 2	4.00	1.00	1.00	.	2.00	1.73
	7	Run 1	1.00	1.00	2.00	1.00	1.25	.50
		Run 2	1.00	1.00	1.00	.	1.00	.00
	8	Run 1	4.00	5.00	4.00	4.00	4.25	.50
		Run 2	4.00	5.00	3.00	.	4.00	1.00

Table A-15: Statistical ANOVA Results: Situation Awareness (SASHA)

Source of Variance	Results ($F_{df\ hypothesis,\ df\ error}$, p-value)							
Item	1	2	3	4	5	6	7	8
REP	$F_{1,2}$=2.14; p=.239	$F_{1,2}$=.63; p=.486	$F_{1,2}$=.07; p=.809	$F_{1,2}$=1.00; p=.391	$F_{1,2}$=2.78; p=.194	$F_{1,2}$= 1.00; p=.423	$F_{1,2}$=.00; p=1.000	$F_{1,2}$=.14 ; p=.754
DD	$F_{3,2}$=.31; p=.816	$F_{3,2}$=.07; p=.972	$F_{3,2}$=.81; p=.565	$F_{3,2}$=1.00; p=.500	$F_{3,2}$=2.48; p=.238	$F_{3,2}$=.51; p=.713	$F_{3,2}$=.28; p=.840	$F_{3,2}$= .07; p=.971
RUN	$F_{1,2}$=.27; p=.638	$F_{1,2}$=1.00; p=.391	$F_{1,2}$=.06; p=.824	$F_{1,2}$=.53; p=.519	$F_{1,2}$=1.00; p=.391	$F_{1,2}$=.25; p=.667	$F_{1,2}$=.75; p=.478	$F_{1,2}$=.75 ; p=.478
Run x Rep	$F_{1,2}$=.27; p=.638	$F_{1,2}$=1.00; p=.391	$F_{1,2}$=.53; p=.519	$F_{1,2}$= .06; p=.824	$F_{1,2}$=1.00 ; p=.391	$F_{1,2}$=1.00; p=.423	$F_{1,2}$=.00; p=1.00	$F_{1,2}$=.00 ; p=1.000
Run x DD	$F_{3,2}$=1.00; p=.500	$F_{3,2}$=1.00; p=.500	$F_{3,2}$=1.16; p=.454	$F_{3,2}$=.53; p=.693	$F_{3,2}$=1.00 ; p=.500	$F_{3,2}$= 4.83; p=.176	$F_{3,2}$=.28; p=.840	$F_{3,2}$=.28 ; p=.840

Note. N=24 Rep=Repetitiveness; DD=Sequence of DD; Int=Interval. xxxp<.001. xxp<.01. xp<.05. (x)p<0.2.

A.2.3. RESULTS FOR PERFORMANCE DATA AND BEHAVIORAL ANALYSIS

Table A-16: Descriptive Statistics: Vienna Reaction Test

a) RT median reaction time

Repetitiveness	Sequence of DD	M	SD	N
non repetitive	l-h	537.00	.	1
	h-l	736.00	.	1
	l-l	527.00	.	1
	h-h	520.00	.	1
	Total	580.00	104.23	4
repetitive	l-h	527.00	.	1
	h-l	586.00	.	1
	l-l	583.00	.	1
	h-h	774.00	.	1
	Total	617.50	107.80	4
Total	l-h	532.00	7.07	2
	h-l	661.00	106.06	2
	l-l	555.00	39.59	2
	h-h	647.00	179.60	2
	Total	598.75	100.19	8

b) RT median motor time

Repetitiveness	Sequence of DD	M	SD	N
non repetitive	l-h	129.00	.	1
	h-l	192.00	.	1
	l-l	249.00	.	1
	h-h	91.00	.	1
	Total	165.25	69.66	4
repetitive	l-h	352.00	.	1
	h-l	208.00	.	1
	l-l	143.00	.	1
	h-h	158.00	.	1
	Total	215.25	95.31	4
Total	l-h	240.50	157.68	2
	h-l	200.00	11.31	2
	l-l	196.00	74.95	2
	h-h	124.50	47.38	2
	Total	190.25	81.77	8

c) RT distribution reaction time

Repetitiveness	Sequence of DD	M	SD	N
non repetitive	l-h	28.49	.	1
	h-l	31.20	.	1
	l-l	26.38	.	1
	h-h	30.00	.	1
	Total	29.02	2.08	4
repetitive	l-h	26.47	.	1
	h-l	27.67	.	1
	l-l	17.25	.	1
	h-h	35.01	.	1
	Total	26.60	7.29	4
Total	l-h	27.48	1.43	2
	h-l	29.44	2.49	2
	l-l	21.82	6.45	2
	h-h	32.50	3.54	2
	Total	27.81	5.13	8

d) RT median motor time

Repetitiveness	Sequence of DD	M	SD	N
non repetitive	l-h	129.00	.	1
	h-l	192.00	.	1
	l-l	249.00	.	1
	h-h	91.00	.	1
	Total	165.25	69.66	4
repetitive	l-h	352.00	.	1
	h-l	208.00	.	1
	l-l	143.00	.	1
	h-h	158.00	.	1
	Total	215.25	95.31	4
Total	l-h	240.50	157.68	2
	h-l	200.00	11.31	2
	l-l	196.00	74.95	2
	h-h	124.50	47.37	2
	Total	190.25	81.77	8

Table A-17: Descriptive Statistics: Cognitrone

a) Sum "correct reactions" COG

Repetitiveness	Sequence of DD	M	SD	N
non repetitive	l-l	86.00	.	1
	h-h	43.00	.	1
	Total	64.50	30.41	2
repetitive	l-h	63.00	.	1
	h-l	64.00	.	1
	h-h	75.00	.	1
	Total	67.33	6.66	3
Total	l-h	63.00	.	1
	h-l	64.00	.	1
	l-l	86.00	.	1
	h-h	59.00	22.63	2
	Total	66.20	15.99	5

b) Sum “incorrect reactions” COG

Repetitiveness	Sequence of DD	M	SD	N
non repetitive	l-l	35.00	.	1
	h-h	19.00	.	1
	Total	27.00	11.31	2
repetitive	l-h	37.00	.	1
	h-l	42.00	.	1
	h-h	65.00	.	1
	Total	48.00	14.93	3
Total	l-h	37.00	.	1
	h-l	42.00	.	1
	l-l	35.00	.	1
	h-h	42.00	32.53	2
	Total	39.60	16.61	5

c) Sum "incorrect non-reactions"

Repetitiveness	Sequence of DD	M	SD	N
non repetitive	l-h	104.00	.	1
	h-l	104.00	.	1
	l-l	18.00	.	1
	h-h	61.00	.	1
	Total	71.75	41.17	4
repetitive	l-h	41.00	.	1
	h-l	40.00	.	1
	h-h	29.00	.	1
	Total	36.67	6.66	3
Total	l-h	72.50	44.55	2
	h-l	72.00	45.26	2
	l-l	18.00	.	1
	h-h	45.00	22.63	2
	Total	56.71	34.84	7

d) Mean time “correct reactions” (sec)

Repetitiveness	Sequence of DD	M	SD	N
non repetitive	l-l	1.25	.	1
	h-h	1.30	.	1
	Total	1.28	.04	2
repetitive	l-h	1.36	.	1
	h-l	1.30	.	1
	h-h	1.29	.	1
	Total	1.32	.04	3
Total	l-h	1.36	.	1
	h-l	1.30	.	1
	l-l	1.250	.	1
	h-h	1.30	.01	2
	Total	1.30	.04	5

e) Mean time "incorrect reactions" (sec)

Repetitiveness	Sequence of DD	M	SD	N
non repetitive	l-l	1.27	.	1
	h-h	1.22	.	1
	Total	1.24	.04	2
repetitive	l-h	1.33	.	1
	h-l	1.23	.	1
	h-h	1.30	.	1
	Total	1.28	.05	3
Total	l-h	1.33	.	1
	h-l	1.23	.	1
	l-l	1.27	.	1
	h-h	1.26	.06	2
	Total	1.27	.05	5

Table A-18: Descriptive Statistics: ZBA

a) Median deviation time total ZBA

Repetitiveness	Sequence of DD	M	SD	N
non repetitive	l-h	.81	.	1
	h-l	1.72	.	1
	l-l	.53	.	1
	h-h	3.06	.	1
	Total	1.53	1.13	4
repetitive	h-l	2.00	.	1
	l-l	.58	.	1
	h-h	.94	.	1
	Total	1.17	.73820	3
Total	l-h	.81	.	1
	h-l	1.86	.19	2
	l-l	.55	.03	2
	h-h	2.00	1.49	2
	Total	1.37	.93	7

b) Median direction deviation total ZBA

Repetitiveness	Sequence of DD	M	SD	N
non repetitive	l-h	82.00	.	1
	h-l	113.00	.	1
	l-l	51.00	.	1
	h-h	41.00	.	1
	Total	71.75	32.57	4
repetitive	h-l	105.00	.	1
	l-l	56.00	.	1
	h-h	119.00	.	1
	Total	93.33	33.08	3
Total	l-h	82.00	.	1
	h-l	109.00	5.65	2
	l-l	53.50	3.53	2
	h-h	80.00	55.15	2
	Total	81.00	32.06	7

A.2.4. Additional Effect Size Calculations

Table A-19: Effect size for each factor

	Non repetitive	repetitive	M	SD	pooled SD	Effect Size DD
Low DD	-0.7	1.33	0.26	1.10		
	-0.68	1.09			0.70	0.78
High DD	-0.32	-0.46	-0.3725	0.30		
	-0.71	0				

M	-0.6025	0.49
SD	0.19	0.86
pooled SD	0.52	
Effect Size Rep	2.08	

Note. z-values of monotony indicator averaged for each condition

Formula 1. Composed indicator of monotony

Indicator of $Monotony_r = HRz_r + Sz_r + FMz_r$

HRz_r = inverted HR summarized for each run (z-score)
Sz_r level-corrected sleepiness-ratings summarized for each run (z-score)
FMz_r ratings of the feeling of monotony after each run (z-score)
$Index_r$ = marks the simulation run

Appendix B: STUDY II

B.1. METHOD

B.1.1. A priory Power Analysis: Additional Information

Based on GPOWER
http://www.psycho.uni-duesseldorf.de/aap/projects/gpower/reference/reference_manual_07.html#t4

a=2 levels (DD;REP)
b=2 levels RUN

REP
Numerator df=2-1=1
Denominator df=24-2=22
Lambda=24 x (((2/1+(2-1)) x0.0625)=1.714
Rho=0.5 assumed

RUN
Numerator df=1
Denominator=(24-2)x(2-1)=22
Lamda=Nxm(0.0625/1-0.5)

B.1.2. SAMPLE DESCRIPTION

Table B-20: Frequencies for Nationalities

Nationality	Frequency	Percent
british	3	12.5
german	6	25.0
dutch	3	12.5
belgian	3	12.5
swedish	2	8.3
danish	2	8.3
finnish	1	4.2
italian	2	8.3
swiss	1	4.2
bulgarian	1	4.2
Total	24	100.0

Table B-21: Statistical Analysis for group differences in age, strain-recovery state and experience

Source	Dependent Variable	df	F	p
rep	age	1	.01	.925
	license	1	.10	.753
	State of Strain	1	1.73	.207
	State of Recovery	1	.05	.829
DD	age	1	1.25	.280
	license	1	.82	.377
	State of Strain	1	1.08	.315
	State of Recovery	1	.12	.734
activity	age	1	.86	.368
	license	1	.19	.669
	State of Strain	1	4.56	.049
	State of Recovery	1	.72	.408
rep x DD	age	1	1.72	.208
	license	1	1.60	.224
	State of Strain	1	.84	.372
	State of Recovery	1	.76	.397
rep x activity	age	1	.98	.336
	license	1	.63	.439
	State of Strain	1	2.93	.106
	State of Recovery	1	2.85	.111
DD x activity	age	1	.03	.875
	license	1	.04	.845
	State of Strain	1	2.75	.117
	State of Recovery	1	.87	.365
rep x DD x activity	age	1	.01	.925
	license	1	.00	.960
	State of Strain	1	.56	.466
	State of Recovery	1	.01	.926
Error	age	16		
	license	16		
	State of Strain	16		
	State of Recovery	16		

B.2. DESCRIPTIVE STATISTICS AND ADDITIONAL RESULTS OF STATISTICAL ANALYSIS

B.2.1. Confirmative Hypotheses

Table B-22: Descriptive statistics for HII.2

a) Initial recovery

				Sequence of DD			
				l-h		h-l	
	Recovery	Repetitiveness		M	SD	M	SD
HR (inv.)	low	non repetitive	Run 1	-8.27	2.18	-4.27	.60
			Run 2	-6.96	1.65	-3.37	.62
		repetitive	Run 1	-3.92	1.55	-4.82	3.43
			Run 2	-.92	2.94	-2.03	2.23
	high	non repetitive	Run 1	-4.56	3.77	-4.27	1.89
			Run 2	-4.49	.92	-2.31	2.69
		repetitive	Run 1	-3.82	.98	-4.47	4.54
			Run 2	-2.90	1.51	-1.29	3.10
Sleepiness	low	non repetitive	Run 1	-.37	.44	-.33	.38
			Run 2	.08	.35	.56	.32
		repetitive	Run 1	.38	.93	-.08	.07
			Run 2	-.18	.42	.58	.54
	high	non repetitive	Run 1	-.19	.38	-.93	.82
			Run 2	-.30	1.01	.62	.39
		repetitive	Run 1	-.18	.60	-.79	.23
			Run 2	.82	.91	1.46	.33
Feeling of monotony	low	non repetitive	Run 1	17	10	43	9
			Run 2	44	39	47	20
		repetitive	Run 1	84	19	61	16
			Run 2	44	29	75	14
	high	non repetitive	Run 1	59	18	46	30
			Run 2	75	9	68	21
		repetitive	Run 1	84	7	49	12
			Run 2	49	18	66	24

b) Initial Strain

				Sequence of DD			
				l-h		h-l	
	Strain	Repetitiveness		M	SD	M	SD
HR (inv.)	low	non repetitive	Run 1	-4.35	5.30	-3.93	.44
			Run 2	-4.47	1.31	-1.18	2.08
		repetitive	Run 1	-3.68	1.30	-4.47	4.54
			Run 2	-.36	2.40	-1.29	3.10
	high	non repetitive	Run 1	-7.45	2.42	-4.44	1.56
			Run 2	-6.35	1.82	-3.67	1.20
		repetitive	Run 1	-4.06	1.26	-4.82	3.43
			Run 2	-3.45	.96	-2.03	2.23
Sleepiness	low	non repetitive	Run 1	-.03	.38	-.93	1.23
			Run 2	-.87	.33	.57	.42
		repetitive	Run 1	.34	.99	-.79	.23
			Run 2	.34	1.32	1.46	.33
	high	non repetitive	Run 1	-.40	.37	-.48	.37
			Run 2	.27	.47	.60	.34
		repetitive	Run 1	-.14	.55	-.08	.07
			Run 2	.30	.10	.58	.54

				Sequence of DD			
				l-h		h-l	
	Strain	Repetititveness		M	SD	M	SD
Feeling of monotony	low	non repetitive	Run 1	52	19	33	27
			Run 2	72	9	48	13
		repetitive	Run 1	89	10	49	12
			Run 2	52	20	66	24
	high	non repetitive	Run 1	31	29	50	17
			Run 2	54	37	62	25
		repetitive	Run 1	78	15	61	16
			Run 2	41	27	75	14

c) Boredom Proneness (BPS)

				Sequence of DD			
				l-h		h-l	
	BPS	Repetititveness		M	SD	M	SD
HR (inv.)	low	non repetitive	Run 1	-5.06	4.50	-4.33	.13
			Run 2	-4.38	.78	-1.72	2.85
		repetitive	Run 1	-4.59	.38	-5.74	4.61
			Run 2	-2.80	1.34	-2.38	2.70
	high	non repetitive	Run 1	-7.77	1.87	-4.24	1.61
			Run 2	-7.06	1.48	-3.40	1.29
		repetitive	Run 1	-2.43	.36	-3.43	3.42
			Run 2	-.13	3.57	-.70	2.80
Sleepiness	low	non repetitive	Run 1	-.04	.40	-1.28	.73
			Run 2	.07	1.03	.88	.02
		repetitive	Run 1	-.06	.82	-.68	.04
			Run 2	.53	.97	1.43	.41
	high	non repetitive	Run 1	-.51	.20	-.31	.35
			Run 2	-.29	.31	.44	.30
		repetitive	Run 1	.42	.78	-.43	.61
			Run 2	-.08	.40	.90	.67
Feeling of monotony	low	non repetitive	Run 1	40	33	25	15
			Run 2	56	42	64	9
		repetitive	Run 1	81	15	45	9
			Run 2	42	22	63	29
	high	non repetitive	Run 1	36	26	54	13
			Run 2	63	24	54	27
		repetitive	Run 1	90	3	62	11
			Run 2	56	27	74	10

Table B-23: Results of Multivariate and Univariate Analysis of Variance HII.2

a) Initial State of Recovery

	Source	Wilks' Lambda	F	p	η_p^2
Between Subjects	rep	.446	5.80	.009	.554
	DD	.829	.96	.437	.171
	recovery	.832	.94	.447	.168
	rep * DD	.847	.84	.494	.153
	rep * recovery	.678	2.22	.131	.322
	DD * recovery	.864	.73	.550	.136
	rep * DD * recovery	.916	.43	.736	.084
Within Subjects	run	.385	7.45	.003	.615
	run * rep	.645	2.57	.096	.355
	run * DD	.556	3.73	.037	.444
	run * recovery	.748	1.57	.241	.252
	run * rep * DD	.637	2.66	.088	.363
	run * rep * recovery	.778	1.33	.304	.222
	run * DD * recovery	.852	.81	.508	.148
	run * rep * DD * recovery	.935	.33	.807	.065

Note. df hypothesis=3, df error=14.

	Source	Measure	F	p	η_p^2
Between Subjects	rep	HR (inv.)	3.54	.078	.18
		sleepiness	8.56	.010	.35
		feeling of monotony	4.52	.049	.22
	DD	HR (inv.)	1.40	.255	.08
		sleepiness	1.06	.319	.06
		feeling of monotony	.00	.990	.00
	recovery	HR (inv.)	.72	.410	.04
		sleepiness	.02	.903	.00
		feeling of monotony	2.34	.146	.13
	rep * DD	HR (inv.)	2.13	.163	.12
		sleepiness	.14	.714	.01
		feeling of monotony	.12	.732	.01
	rep * recovery	HR (inv.)	1.12	.306	.07
		sleepiness	1.87	.190	.10
		feeling of monotony	4.47	.050	.22
	DD * recovery	HR (inv.)	.08	.783	.00
		sleepiness	.39	.543	.02
		feeling of monotony	2.01	.175	.11
	rep * DD * recovery	HR (inv.)	1.13	.304	.07
		sleepiness	.00	.956	.00
		feeling of monotony	.19	.669	.01
Within Subjects	run	HR (inv.)	16.37	.001	.51
		sleepiness	14.50	.002	.48
		feeling of monotony	.37	.550	.02
	run * rep	HR (inv.)	2.60	.126	.14
		sleepiness	.13	.722	.01
		feeling of monotony	7.45	.015	.32
	run * DD	HR (inv.)	1.02	.329	.06
		sleepiness	8.08	.012	.34
		feeling of monotony	4.40	.052	.22
	run * recovery	HR (inv.)	.29	.598	.02
		sleepiness	4.06	.061	.20
		feeling of monotony	.14	.715	.01
	run * rep * DD	HR (inv.)	.03	.872	.00
		sleepiness	.05	.826	.00
		feeling of monotony	8.67	.010	.35
	run * rep * recovery	HR (inv.)	.19	.670	.01
		sleepiness	3.53	.079	.18
		feeling of monotony	.00	.978	.00
	run * DD * recovery	HR (inv.)	1.86	.191	.10

		sleepiness	.60	.449	.04
		feeling of monotony	.41	.529	.03
	run * rep * DD * recovery	HR (inv.)	.00	.959	.00
		sleepiness	.55	.469	.03
		feeling of monotony	.56	.466	.03

Note. df hypothesis=1, df error=16.

b) Initial State of Strain

	Source	Wilks' Lambda	F	p	η_p^2
Between Subjects	rep	.432	6.12	.007	.568
	DD	.812	1.08	.390	.188
	strain	.861	.75	.539	.139
	rep * DD	.860	.76	.534	.140
	rep * strain	.789	1.25	.329	.211
	DD * strain	.825	.99	.427	.175
	rep * DD * strain	.955	.22	.881	.045
Within Subjects	run	.385	7.72	.003	.623
	run * rep	.656	2.45	.107	.344
	run * DD	.537	4.02	.029	.463
	run * strain	.922	.39	.760	.078
	run * rep * DD	.676	2.24	.128	.324
	run * rep * strain	.893	.56	.649	.107
	run * DD * strain	.754	1.52	.252	.246
	run * rep * DD * strain	.849	.83	.501	.151

Note. df hypothesis=3, df error=14.

	Source	Measure	F	p	η_p^2
Between Subjects	rep	HR (inv.)	2.20	.157	.12
		sleepiness	12.18	.003	.43
		feeling of monotony	3.15	.095	.16
	DD	HR (inv.)	1.13	.304	.07
		sleepiness	1.37	.259	.08
		feeling of monotony	.15	.706	.01
	strain	HR (inv.)	2.55	.130	.14
		sleepiness	.34	.568	.02
		feeling of monotony	.01	.909	.00
	rep * DD	HR (inv.)	1.77	.201	.10
		sleepiness	.24	.632	.01
		feeling of monotony	.01	.940	.00
	rep * strain	HR (inv.)	.19	.668	.01
		sleepiness	4.27	.055	.21
		feeling of monotony	.01	.917	.00
	DD * strain	HR (inv.)	.31	.587	.02
		sleepiness	.01	.931	.00
		feeling of monotony	3.21	.092	.17
	rep * DD * strain	HR (inv.)	.00	.959	.00
		sleepiness	.47	.501	.03
		feeling of monotony	.20	.660	.01
Within Subjects	run	HR (inv.)	17.94	.001	.53
		sleepiness	11.20	.004	.41
		feeling of monotony	.33	.576	.02
	run * rep	HR (inv.)	2.52	.132	.14
		sleepiness	.30	.592	.02
		feeling of monotony	6.55	.021	.29
	run * DD	HR (inv.)	1.83	.195	.10
		sleepiness	9.15	.008	.36

		feeling of monotony	4.03	.062	.20
	run * strain	HR (inv.)	1.29	.274	.07
		sleepiness	.00	.975	.00
		feeling of monotony	.00	.976	.00
	run * rep * DD	HR (inv.)	.02	.886	.00
		sleepiness	.03	.874	.00
		feeling of monotony	7.44	.015	.32
	run * rep * strain	HR (inv.)	.48	.499	.03
		sleepiness	1.66	.216	.09
		feeling of monotony	.00	.959	.00
	run * DD * strain	HR (inv.)	.07	.799	.00
		sleepiness	5.22	.036	.25
		feeling of monotony	.03	.856	.00
	run * rep * DD * strain	HR (inv.)	2.65	.123	.14
		sleepiness	.00	.949	.00
		feeling of monotony	.00	.968	.00

Note. df hypothesis=1, df error=16.

c) Boredom Proneness (BPS)

	Source	Wilks' Lambda	F	p	η_p^2
Between Subjects	rep	.343	8.94	.001	.657
	DD	.779	1.32	.307	.221
	bps	.894	.55	.654	.106
	rep * DD	.791	1.24	.334	.209
	rep * bps	.738	1.65	.222	.262
	DD * bps	.841	.88	.473	.159
	rep * DD * bps	.758	1.49	.260	.242
Within Subjects	run	.322	9.84	.001	.678
	run * rep	.601	3.09	.061	.399
	run * DD	.429	6.22	.007	.571
	run * bps	.790	1.24	.332	.210
	run * rep * DD	.647	2.54	.098	.353
	run * rep * bps	.940	.30	.826	.060
	run * DD * bps	.783	1.30	.315	.217
	run * rep * DD * bps	.874	.67	.583	.126

Note. df hypothesis=3, df error=14.

	Source	Measure	F	p	η_p^2
	rep	HR (inv.)	4.66	.046	.23
		sleepiness	10.83	.005	.40
		feeling of monotony	3.50	.080	.18
	DD	HR (inv.)	1.28	.274	.07
		sleepiness	1.00	.331	.06
		feeling of monotony	.12	.736	.01
	bps	HR (inv.)	.06	.803	.00
		sleepiness	.58	.457	.04
		feeling of monotony	1.31	.270	.08
	rep * DD	HR (inv.)	3.12	.097	.16
		sleepiness	.01	.925	.00
		feeling of monotony	.17	.687	.01
	rep * bps	HR (inv.)	4.67	.046	.23
		sleepiness	.02	.888	.00
		feeling of monotony	.19	.665	.01
	DD * bps	HR (inv.)	.17	.689	.01
		sleepiness	1.66	.216	.09
		feeling of monotony	.12	.732	.01

Source	Measure	F	p	η_p^2
rep * DD * bps	HR (inv.)	.40	.534	.02
	sleepiness	2.63	.124	.14
	feeling of monotony	.04	.836	.00
run	HR (inv.)	16.64	.001	.51
	sleepiness	14.37	.002	.47
	feeling of monotony	1.11	.308	.06
run * rep	HR (inv.)	2.11	.166	.12
	sleepiness	.02	.878	.00
	feeling of monotony	10.18	.006	.39
run * DD	HR (inv.)	1.22	.286	.07
	sleepiness	11.05	.004	.41
	feeling of monotony	6.54	.021	.29
run * bps	HR (inv.)	.25	.622	.02
	sleepiness	3.14	.096	.16
	feeling of monotony	.58	.458	.03
run * rep * DD	HR (inv.)	.00	.988	.00
	sleepiness	.19	.669	.01
	feeling of monotony	7.83	.013	.33
run * rep * bps	HR (inv.)	.19	.666	.01
	sleepiness	.10	.760	.01
	feeling of monotony	.50	.488	.03
run * DD * bps	HR (inv.)	.64	.435	.04
	sleepiness	.47	.504	.03
	feeling of monotony	2.40	.141	.13
run * rep * DD * bps	HR (inv.)	.03	.865	.00
	sleepiness	1.05	.320	.06
	feeling of monotony	.97	.340	.06

Note. df hypothesis=1, df error=16.

Table B-24: Descriptive Statistics HII.3

	Repetitiveness	Sequence of DD	Activity	M	SD	n
HR inv. during run 3 (corr.)	non repetitive	l-h	non active	-4.20	.86	3
			active	-3.38	3.15	3
		h-l	non active	.02	2.63	3
			active	-2.85	2.94	3
	repetitive	l-h	non active	1.10	2.49	3
			active	-1.52	3.01	3
		h-l	non active	.31	1.17	3
			active	1.36	1.94	3
Sleepiness before and after run 3 (corr.)	non repetitive	l-h	non active	2.04	.26	3
			active	.25	.66	3
		h-l	non active	1.79	1.63	3
			active	-.04	.56	3
	repetitive	l-h	non active	.33	.63	3
			active	-.92	.75	3
		h-l	non active	-.13	.76	3
			active	-.92	.56	3
Feeling of monotony after run 3	non repetitive	l-h	non active	76.00	3.61	3
			active	56.67	33.65	3
		h-l	non active	71.33	29.77	3
			active	61.00	14.93	3
	repetitive	l-h	non active	65.33	18.01	3
			active	51.67	37.53	3
		h-l	non active	41.67	18.90	3
			active	47.33	33.38	3

Table B-25: Results of statistical analysis HII.3

Source	Dependent Variable	F	df	p
activity	HR inv. during run 3 (corr.)	.84	1	.373
	Sleepiness (corr. run 3)	18.08	1	.001
	Feeling of monotony (run 3)	.78	1	.390
rep	HR inv. during run 3 (corr.)	8.76	1	.009
	Sleepiness (corr. run 3)	18.08	1	.001
	Feeling of monotony (run 3)	1.91	1	.186
DD	HR inv. during run 3 (corr.)	3.01	1	.102
	Sleepiness (corr. run 3)	.56	1	.464
	Feeling of monotony (run 3)	.44	1	.516
rep x activity	HR inv. during run 3 (corr.)	.02	1	.903
	Sleepiness (corr. run 3)	1.41	1	.252
	Feeling of monotony (run 3)	.26	1	.618
DD x activity	HR inv. during run 3 (corr.)	.00	1	.996
	Sleepiness (corr. run 3)	.10	1	.759
	Feeling of monotony (run 3)	.44	1	.516
rep x DD	HR inv. during run 3 (corr.)	.46	1	.508
	Sleepiness (corr. run 3)	.00	1	.951
	Feeling of monotony (run 3)	.42	1	.526
rep x DD x activity	HR inv. during run 3 (corr.)	3.49	1	.080
	Sleepiness (corr. run 3)	.14	1	.712
	Feeling of monotony (run 3)	.06	1	.812
Error	HR inv. during run 3 (corr.)		16	
	Sleepiness (corr. run 3)		16	
	Feeling of monotony (run 3)		16	

B.2.2. ADDITIONAL RESULTS: PHYSIOLOGICAL DATA

Table B-26: Descriptive Statistics for HR (corr.)

Run1			Repetitiveness			
			non repetitive		repetitive	
			M	SD	M	SD
Sequence of DD	l-h	1	6.16	4.77	3.79	2.02
		2	6.39	3.79	4.15	2.13
		3	4.85	4.03	2.99	2.27
		4	6.05	2.28	4.19	1.45
		5	5.45	3.99	4.63	.55
		6	6.64	3.82	4.61	1.30
		7	5.89	3.68	3.70	1.80
		8	7.30	3.34	4.66	2.16
		9	6.93	4.11	3.98	2.05
		10	6.76	4.25	3.73	3.00
		11	6.53	4.93	3.89	3.25
		12	7.27	3.55	4.07	2.34
		13	6.61	3.81	2.90	2.52
		14	7.14	3.64	3.85	2.93
		15	6.27	2.65	2.86	2.44
	h-l	1	2.69	1.96	5.36	2.89
		2	4.00	1.02	4.88	3.29
		3	3.73	1.72	4.43	3.41
		4	4.44	2.35	4.36	3.90
		5	4.03	2.06	4.35	3.38
		6	4.78	1.79	4.37	3.91
		7	4.76	2.52	3.99	4.45
		8	5.11	2.48	3.45	4.74
		9	5.57	2.80	4.40	3.85
		10	5.39	2.33	5.09	4.46
		11	3.45	2.45	5.49	5.37
		12	4.53	2.01	4.61	3.61
		13	3.35	1.85	4.59	3.86
		14	3.56	2.18	3.76	4.55
		15	4.64	2.25	5.67	4.86

Run 2			Repetitiveness			
			non repetitive		repetitive	
			M	SD	M	SD
Sequence of DD	l-h	1	5.34	2.53	1.84	2.68
		2	6.70	2.11	1.47	3.05
		3	5.10	1.97	1.49	2.77
		4	5.48	2.84	1.93	2.49
		5	5.03	2.41	1.89	2.63
		6	6.22	2.54	3.07	3.27
		7	6.34	2.84	1.59	2.85
		8	4.47	2.52	1.36	2.26
		9	5.87	2.20	1.25	3.26
		10	6.74	1.46	.93	2.71
		11	6.15	.82	2.05	2.21
		12	6.34	2.74	1.42	3.60
		13	5.41	1.55	2.49	3.07
		14	5.28	1.02	2.10	2.72
		15	5.35	2.71	2.07	2.36
	h-l	1	2.08	2.18	2.65	2.70
		2	2.31	2.16	3.35	3.50
		3	2.34	1.30	1.78	2.76
		4	2.69	2.05	1.55	2.79
		5	2.47	2.61	1.62	3.29
		6	3.05	2.06	2.08	2.56
		7	2.83	1.65	1.80	3.67
		8	3.85	3.09	1.85	3.76
		9	3.00	1.60	.87	3.13
		10	3.96	2.31	1.29	3.26
		11	2.78	3.95	1.12	1.82
		12	2.40	2.25	.03	1.79
		13	3.02	2.58	.80	3.46
		14	2.26	2.13	1.30	3.90
		15	3.54	3.05	1.00	3.05

Table B-27: Descriptive Statistics for HRV

HRV Run 1			non repetitive		repetitive	
			M	SD	M	SD
Sequence of DD	l-h	1	27.07	12.88	26.21	14.48
		2	19.33	6.46	28.76	14.93
		3	15.09	3.78	18.09	10.58
		4	19.73	8.38	20.10	9.22
		5	17.23	7.65	35.32	18.62
		6	17.38	5.00	44.10	28.16
		7	19.23	7.69	34.41	22.93
		8	15.83	3.87	41.42	17.35
		9	16.94	10.03	31.32	13.44
		10	26.07	12.35	46.06	33.61
		11	21.15	7.23	40.22	21.33
		12	20.91	11.90	62.52	35.08
		13	20.29	10.73	38.30	11.36
		14	22.77	7.55	28.79	17.20
		15	22.70	9.89	30.03	12.77
	h-l	1	25.79	12.41	33.89	24.12
		2	19.80	8.98	25.59	13.73
		3	21.04	14.94	22.29	10.96
		4	28.26	18.48	25.45	15.68
		5	21.89	11.43	31.09	13.01
		6	25.22	13.90	45.34	16.59
		7	23.58	15.81	29.02	19.60
		8	19.22	12.81	26.99	12.11
		9	27.01	15.69	28.64	15.73
		10	23.03	12.37	22.24	11.72
		11	20.65	15.50	24.32	4.94
		12	30.01	19.90	32.96	16.04
		13	22.18	9.05	31.13	15.00
		14	24.21	16.03	22.58	11.90
		15	34.95	26.04	28.85	14.26

Run 2			Repetitiveness			
			non repetitive		repetitive	
			M	SD	M	SD
Sequence of DD	l-h	1	29.98	13.56	47.73	54.76
		2	25.18	11.81	54.51	44.84
		3	22.29	10.11	47.01	20.68
		4	21.79	9.53	31.82	13.04
		5	25.60	16.21	38.39	23.16
		6	20.57	6.44	66.56	43.43
		7	24.44	11.52	47.40	25.56
		8	23.57	7.99	36.82	27.17
		9	25.85	10.06	68.80	62.09
		10	29.31	4.91	39.67	10.33
		11	24.85	11.40	46.95	32.36
		12	25.07	8.51	30.59	15.40
		13	31.90	10.85	32.02	18.95
		14	20.08	11.24	34.70	11.91
		15	25.85	10.95	27.82	9.65
	h-l	1	22.44	14.82	46.49	27.85
		2	20.46	12.72	41.49	19.98
		3	25.52	21.07	35.86	19.45
		4	34.79	24.06	36.15	14.34
		5	30.36	16.87	29.16	13.54
		6	28.72	13.62	34.17	9.63
		7	39.61	33.65	31.73	10.09
		8	28.17	21.54	37.95	12.72
		9	29.07	25.02	38.33	15.68
		10	35.80	17.13	50.43	9.50
		11	34.08	22.21	40.59	17.89
		12	21.35	9.01	36.08	7.24
		13	28.96	17.92	42.42	14.00
		14	21.05	13.33	33.55	19.05
		15	27.31	9.21	31.35	22.35

Table B-28: Descriptive Statistics for No. of Blinks

Run 1			Repetitiveness: non repetitive M	non repetitive SD	repetitive M	repetitive SD
Sequence of DD	l-h	1	30.83	18.95	22.17	15.11
		2	32.33	21.40	26.17	18.77
		3	32.83	20.87	26.33	22.75
		4	36.83	20.65	28.50	22.75
		5	37.17	19.71	24.33	16.97
		6	32.50	21.55	24.50	13.82
		7	34.67	17.31	24.83	19.27
		8	35.17	22.92	30.00	27.18
		9	37.67	20.89	26.50	19.39
		10	39.67	23.82	27.00	20.74
		11	38.83	23.75	24.17	19.14
		12	40.83	22.81	29.50	24.72
		13	42.83	23.03	22.83	20.66
		14	48.17	18.10	22.67	20.11
		15	45.50	19.95	27.00	21.31
	h-l	1	14.17	7.08	14.00	8.27
		2	14.33	6.80	21.33	12.39
		3	15.33	9.69	16.00	9.14
		4	18.17	8.95	17.83	9.91
		5	18.00	10.39	23.50	13.22
		6	20.17	10.34	17.83	9.06
		7	17.33	9.11	20.17	9.75
		8	17.50	10.91	22.67	12.11
		9	17.80	8.70	24.50	7.42
		10	17.60	8.41	19.33	7.53
		11	17.50	14.20	17.00	9.70
		12	18.83	8.50	17.83	8.57
		13	16.80	10.38	25.00	18.03
		14	26.17	20.83	22.67	13.59
		15	29.00	22.61	24.67	15.62

Run 2			Repetitiveness: non repetitive M	non repetitive SD	repetitive M	repetitive SD
Sequence of DD	l-h	1	38.50	27.54	18.17	9.37
		2	42.80	20.47	22.17	15.09
		3	43.00	29.48	29.17	28.12
		4	47.80	18.81	23.83	19.78
		5	42.40	20.51	28.33	25.10
		6	42.17	26.41	23.33	14.54
		7	44.50	22.69	27.67	18.72
		8	45.67	22.51	29.50	22.76
		9	43.67	22.93	25.50	19.43
		10	44.17	26.35	20.67	11.08
		11	41.00	26.94	20.33	15.44
		12	41.33	25.47	27.00	23.38
		13	51.00	29.09	27.67	22.65
		14	46.83	22.92	24.33	21.10
		15	48.17	24.08	17.40	13.79
	h-l	1	15.00	13.16	26.00	10.60
		2	16.33	9.48	31.00	17.34
		3	18.50	10.75	31.17	15.48
		4	25.17	14.52	26.00	11.12
		5	22.33	13.32	29.50	20.77
		6	20.50	10.73	27.33	19.41
		7	23.83	12.73	35.00	21.73
		8	21.00	11.03	35.33	24.90
		9	23.33	13.08	39.33	29.54
		10	24.17	12.38	34.50	18.57
		11	21.00	9.76	25.33	20.51
		12	16.83	8.82	31.17	23.15
		13	19.67	7.97	32.67	22.77
		14	22.67	13.41	29.00	20.12
		15	24.83	16.12	25.33	15.11

Table B-29: Descriptive Statistics for SCL (corr.)

Run 1			Repetitiveness			
			non repetitive		repetitive	
			M	SD	M	SD
Sequence of DD	l-h	1	.36	.29	.29	.17
		2	.34	.26	.29	.18
		3	.34	.26	.29	.18
		4	.34	.26	.29	.18
		5	.33	.26	.28	.19
		6	.39	.27	.36	.13
		7	.34	.24	.35	.14
		8	.34	.24	.35	.14
		9	.29	.21	.34	.18
		10	.29	.21	.34	.17
		11	.34	.21	.32	.16
		12	.33	.17	.32	.18
		13	.31	.18	.35	.15
		14	.31	.18	.35	.15
		15	.31	.18	.35	.15
	h-l	1	.22	.12	.21	.09
		2	.21	.11	.20	.08
		3	.21	.11	.20	.08
		4	.21	.11	.20	.08
		5	.20	.10	.19	.05
		6	.29	.16	.32	.08
		7	.26	.14	.27	.07
		8	.26	.14	.27	.07
		9	.23	.14	.23	.08
		10	.25	.13	.24	.10
		11	.25	.15	.26	.07
		12	.19	.13	.18	.06
		13	.26	.15	.19	.05
		14	.26	.15	.19	.05
		15	.26	.15	.19	.05

Run 2			Repetitiveness			
			non repetitive		repetitive	
			M	SD	M	SD
Sequence of DD	l-h	1	.50	.29	.43	.11
		2	.39	.28	.36	.15
		3	.32	.28	.36	.20
		4	.33	.37	.34	.14
		5	.35	.32	.43	.19
		6	.33	.26	.39	.07
		7	.29	.24	.42	.25
		8	.36	.27	.38	.20
		9	.33	.27	.38	.21
		10	.33	.27	.38	.21
		11	.29	.27	.38	.22
		12	.29	.27	.38	.22
		13	.29	.30	.23	.07
		14	.30	.32	.24	.10
		15	.32	.36	.24	.14
	h-l	1	.46	.34	.41	.17
		2	.30	.14	.29	.12
		3	.26	.13	.28	.10
		4	.24	.17	.23	.10
		5	.23	.20	.18	.12
		6	.34	.20	.23	.15
		7	.30	.18	.14	.17
		8	.26	.13	.13	.18
		9	.25	.14	.22	.14
		10	.25	.14	.22	.14
		11	.24	.16	.24	.12
		12	.24	.16	.24	.12
		13	.24	.16	.25	.09
		14	.24	.16	.28	.12
		15	.33	.25	.30	.18

Table B-30: Descriptive Statistics for Theta (corr.)

Run 1			Repetitiveness			
			non repetitive		repetitive	
			M	SD	M	SD
Sequence of DD	l-h	1	-2.39	3.25	-1.18	1.49
		2	-1.60	2.11	-.39	1.40
		3	-1.93	2.14	-.62	1.34
		4	-1.72	2.36	-.82	1.61
		5	-1.94	2.14	-1.10	1.96
		6	-2.32	2.92	-1.18	1.94
		7	-2.10	2.36	-1.29	2.04
		8	-1.96	2.22	-1.10	1.99
		9	-2.01	2.50	-1.00	1.89
		10	-2.08	2.46	-.51	1.92
		11	-2.32	3.11	-.92	1.74
		12	-1.50	1.91	-.54	1.04
		13	-1.96	2.19	-1.07	2.06
		14	-1.48	1.56	-1.02	1.80
		15	-2.13	2.43	-.87	1.52
	h-l	1	-1.66	1.60	-.50	2.03
		2	-1.52	2.02	-.01	2.37
		3	-2.08	1.35	-.26	2.06
		4	-1.59	1.83	-.35	1.93
		5	-1.60	1.56	-.50	1.67
		6	-1.68	1.42	-.24	2.09
		7	-2.02	1.43	-.67	1.63
		8	-1.72	1.24	-.27	2.19
		9	-1.77	1.60	-.76	1.55
		10	-1.56	1.90	-.63	1.82
		11	-1.92	1.13	-.29	2.38
		12	-1.72	1.25	-.59	2.04
		13	-1.84	1.18	-.73	1.85
		14	-1.77	1.39	-.57	2.01
		15	-2.01	1.36	-.36	2.57

Run 2			Repetitiveness			
			non repetitive		repetitive	
			M	SD	M	SD
Sequence of DD	l-h	1	-1.75	1.66	.06	.99
		2	-1.90	2.53	.30	1.28
		3	-1.80	1.87	-1.05	2.24
		4	-1.99	2.28	-.35	1.85
		5	-1.47	1.25	-1.04	2.65
		6	-2.17	2.35	-.68	2.16
		7	-2.34	2.63	-.76	1.84
		8	-1.63	1.49	-.82	2.10
		9	-2.12	1.83	-.60	1.56
		10	-2.12	2.15	-.60	1.79
		11	-2.23	2.37	-.25	1.40
		12	-1.46	1.37	-.71	1.81
		13	-.90	1.34	-.76	1.93
		14	-1.66	1.40	-.67	1.70
		15	-2.23	2.33	-1.01	2.37
	h-l	1	-1.42	1.63	-.43	2.09
		2	-1.02	2.24	-.15	2.29
		3	-1.32	1.89	.15	2.56
		4	-1.34	1.64	-.74	1.75
		5	-1.44	1.79	-.33	2.24
		6	-1.31	2.08	-.20	2.14
		7	-1.00	1.75	-.23	1.93
		8	-1.14	1.70	-.43	1.80
		9	-1.27	1.61	-.13	1.92
		10	-.69	2.21	-.32	2.43
		11	-1.16	2.03	.12	2.19
		12	-.68	2.45	.05	2.17
		13	-1.03	2.10	-.11	1.91
		14	-1.44	1.90	.11	2.77
		15	-1.80	1.60	-.33	2.19

Table B-31: Descriptive Statistics for Alpha1 (corr.)

			Repetitiveness			
			non repetitive		repetitive	
Run 1			M	SD	M	SD
Sequence of DD	l-h	1	-14.72	18.46	-12.53	8.97
		2	-14.11	17.49	-11.86	8.15
		3	-14.58	17.40	-11.95	8.12
		4	-13.72	16.73	-11.97	8.37
		5	-13.88	16.94	-11.85	8.66
		6	-14.26	17.91	-11.85	8.82
		7	-13.70	16.87	-11.72	8.51
		8	-13.79	16.88	-11.81	8.78
		9	-14.22	17.31	-11.53	8.36
		10	-14.15	17.23	-11.13	8.73
		11	-14.45	17.72	-11.66	8.00
		12	-13.97	17.04	-11.29	8.01
		13	-13.75	16.79	-11.58	8.10
		14	-13.53	16.76	-12.18	8.12
		15	-13.07	16.07	-11.51	8.23
	h-l	1	-14.85	11.94	-13.60	13.56
		2	-14.74	12.08	-13.13	13.15
		3	-14.69	11.82	-12.99	12.89
		4	-14.83	11.79	-13.07	12.67
		5	-14.59	11.97	-13.30	13.47
		6	-14.82	11.71	-12.94	13.32
		7	-14.53	11.52	-12.91	12.92
		8	-14.79	11.75	-13.11	13.13
		9	-14.50	11.49	-12.71	12.45
		10	-14.90	11.42	-12.86	12.60
		11	-15.08	11.66	-13.19	13.00
		12	-14.64	11.68	-12.98	12.65
		13	-14.75	12.03	-12.70	12.75
		14	-14.63	11.61	-13.05	13.06
		15	-14.45	11.54	-12.98	13.27

			Repetitiveness			
			non repetitive		repetitive	
Run 2			M	SD	M	SD
Sequence of DD	l-h	1	-13.44	16.37	-11.03	7.89
		2	-13.94	17.24	-10.50	7.07
		3	-13.44	16.32	-11.07	7.51
		4	-14.34	17.46	-11.10	8.27
		5	-13.83	17.13	-11.40	8.95
		6	-14.16	17.36	-11.31	8.57
		7	-13.87	16.62	-10.46	7.26
		8	-13.97	16.93	-11.25	8.44
		9	-13.25	15.93	-10.70	8.39
		10	-13.67	16.82	-10.90	7.56
		11	-14.27	17.47	-11.32	8.78
		12	-12.63	15.52	-10.71	7.78
		13	-11.45	14.12	-10.20	7.18
		14	-13.58	16.54	-11.42	8.43
		15	-13.77	17.13	-11.60	8.59
	h-l	1	-14.11	10.58	-12.60	12.43
		2	-13.78	10.67	-12.34	12.46
		3	-13.63	10.32	-12.47	12.52
		4	-13.68	10.63	-12.47	12.54
		5	-14.00	10.80	-12.71	12.90
		6	-14.41	10.58	-12.51	12.96
		7	-14.05	10.64	-12.43	13.01
		8	-13.84	10.95	-12.30	12.44
		9	-13.64	9.89	-12.06	12.64
		10	-13.61	10.66	-12.48	12.39
		11	-13.78	10.97	-12.27	12.27
		12	-14.05	10.71	-11.80	12.08
		13	-13.60	10.66	-12.21	12.35
		14	-13.54	10.87	-12.39	12.53
		15	-13.89	10.72	-12.70	12.97

Table B-32: Descriptive Statistics for Alpha2 (corr.)

Run 1			Repetitiveness: non repetitive M	non repetitive SD	repetitive M	repetitive SD
Sequence of DD	l-h	1	-4.80	.84	-7.79	5.37
		2	-4.72	1.08	-7.65	5.10
		3	-4.94	1.17	-7.71	5.00
		4	-3.87	1.56	-7.53	4.98
		5	-4.36	1.15	-7.49	4.76
		6	-4.62	1.52	-7.51	4.55
		7	-4.55	1.21	-6.90	4.49
		8	-4.68	1.20	-6.89	4.11
		9	-4.64	1.02	-7.23	4.66
		10	-4.39	1.45	-7.05	4.34
		11	-4.50	.77	-7.60	4.59
		12	-4.39	.97	-7.29	4.46
		13	-4.60	.84	-7.82	5.00
		14	-4.20	1.05	-7.70	4.88
		15	-3.51	1.24	-7.42	4.73
	h-l	1	-7.54	5.57	-3.74	3.45
		2	-7.51	5.63	-3.78	3.75
		3	-7.47	5.91	-3.78	3.71
		4	-7.07	5.95	-3.75	3.70
		5	-7.31	5.87	-3.50	3.72
		6	-7.43	5.53	-3.41	3.59
		7	-7.25	5.66	-3.57	3.76
		8	-7.67	5.81	-3.58	3.76
		9	-7.24	5.83	-3.55	3.68
		10	-7.56	5.74	-3.76	3.69
		11	-7.86	5.59	-3.72	3.54
		12	-7.53	5.76	-3.56	3.75
		13	-7.77	5.72	-3.53	3.70
		14	-7.79	5.58	-3.62	3.70
		15	-7.22	5.95	-3.64	3.70

Run 2			Repetitiveness: non repetitive M	non repetitive SD	repetitive M	repetitive SD
Sequence of DD	l-h	1	-4.15	1.48	-7.08	4.59
		2	-4.70	.83	-7.22	4.63
		3	-4.41	1.23	-7.27	4.50
		4	-4.60	1.03	-7.44	4.66
		5	-4.76	.90	-7.53	4.61
		6	-4.53	.82	-7.49	4.96
		7	-4.84	.79	-7.04	4.43
		8	-4.53	.76	-7.10	4.26
		9	-4.32	1.57	-6.96	4.18
		10	-4.87	.63	-7.44	4.50
		11	-4.32	.77	-7.67	4.60
		12	-4.49	1.36	-6.74	4.08
		13	-3.65	2.41	-6.92	4.55
		14	-4.50	1.28	-7.23	4.56
		15	-3.79	1.72	-6.81	4.25
	h-l	1	-6.23	6.04	-3.49	3.87
		2	-6.52	5.84	-3.59	3.89
		3	-6.73	5.79	-3.45	3.77
		4	-6.32	6.05	-3.48	3.73
		5	-6.75	5.90	-3.23	3.69
		6	-7.17	5.60	-3.26	3.50
		7	-6.53	5.78	-3.26	3.70
		8	-6.20	6.36	-3.13	3.67
		9	-6.40	5.62	-3.23	3.54
		10	-6.28	5.63	-3.34	3.68
		11	-6.34	5.68	-3.28	3.61
		12	-6.75	5.48	-3.28	3.53
		13	-6.41	5.55	-3.29	3.48
		14	-6.24	6.11	-3.76	3.98
		15	-6.79	5.64	-3.59	3.58

Table B-33: Descriptive Statistics for Beta (corr.)

Run 1			Repetitiveness: non repetitive M	non repetitive SD	repetitive M	repetitive SD
Sequence of DD	l-h	1	.21	3.80	.13	4.16
		2	2.10	7.38	.40	5.48
		3	3.52	7.60	-.94	4.32
		4	7.60	15.58	-.81	3.68
		5	10.88	26.05	-.05	3.85
		6	13.64	31.57	.34	3.35
		7	1.56	6.01	.44	4.18
		8	-.17	1.52	.99	4.16
		9	.78	4.26	.37	3.96
		10	4.72	11.11	-.02	3.61
		11	5.13	9.75	-.20	2.81
		12	2.01	7.09	.42	3.83
		13	.25	3.40	.47	4.30
		14	5.17	12.74	1.07	3.63
		15	4.84	7.84	.76	3.32
	h-l	1	.30	5.26	1.76	3.13
		2	.48	5.27	1.28	3.96
		3	-.06	5.50	.60	5.19
		4	.12	5.65	.81	5.22
		5	-1.71	5.46	2.22	4.61
		6	2.91	6.61	5.10	5.42
		7	2.64	9.04	.12	6.99
		8	-.83	5.75	-1.07	5.49
		9	-.37	6.02	-1.47	4.61
		10	-1.19	5.76	-1.38	4.32
		11	-1.04	5.24	2.35	5.50
		12	-1.11	5.46	1.53	3.58
		13	-1.84	5.17	.12	5.22
		14	-1.92	4.84	1.17	6.29
		15	-.36	4.18	1.74	5.40

Run 2			Repetitiveness: non repetitive M	non repetitive SD	repetitive M	repetitive SD
Sequence of DD	l-h	1	-3.08	3.68	-4.97	2.49
		2	-5.20	1.96	-5.25	2.99
		3	-4.78	1.36	-5.32	2.45
		4	-4.38	1.46	-5.37	2.28
		5	-4.57	1.66	-4.98	2.16
		6	-3.88	2.19	-4.86	1.90
		7	-4.83	1.54	-4.89	2.49
		8	-4.27	1.94	-5.17	2.52
		9	-4.61	1.84	-4.69	2.24
		10	-4.75	1.57	-4.99	2.46
		11	-3.21	3.44	-5.07	2.08
		12	-4.03	2.41	-4.88	2.14
		13	-3.89	2.54	-4.32	1.99
		14	-3.74	2.78	-4.46	1.75
		15	-3.99	1.72	-4.38	1.87
	h-l	1	-6.07	4.40	-5.31	5.65
		2	-6.25	3.89	-5.61	5.87
		3	-5.86	3.37	-5.40	5.67
		4	-4.02	5.72	-5.11	5.43
		5	-6.51	3.48	-4.21	3.86
		6	-6.70	3.80	-3.40	3.88
		7	-5.44	3.89	-4.98	5.39
		8	-3.80	6.75	-4.96	5.85
		9	-6.58	3.43	-5.14	5.40
		10	-5.88	3.24	-4.82	4.33
		11	-5.52	3.62	-3.48	3.69
		12	-6.43	3.06	-4.03	4.03
		13	-5.28	3.97	-4.84	4.50
		14	-5.15	4.43	-5.87	5.88
		15	-6.33	3.34	-5.99	5.58

B.2.3. ADDITIONAL RESULTS: SUBJECTIVE DATA

Table B-34: Descriptive Statistics for Thackray Items

attentiveness		Repetitiveness			
		non repetitive		repetitive	
		M	SD	M	SD
Sequence of DD	l-h	1.35	.63	-.30	.34
		.18	.91	-.30	.34
		-.65	.66	-.13	.49
		.18	1.23	.20	.24
		-.15	.80	.20	.24
		-.82	.68	-.13	.75
	h-l	1.03	.61	.30	.35
		.20	1.02	.47	.36
		-.13	.87	.30	.35
		.53	.52	-.20	.38
		-.13	.69	-.70	.64
		-.80	.67	-.37	.52

boredom		Repetitiveness			
		non repetitive		repetitive	
		M	SD	M	SD
Sequence of DD	l-h	-.85	1.48	.33	1.07
		-.68	1.57	1.17	1.36
		.98	.95	.00	.75
		.32	1.09	-.33	.74
		.15	1.10	-.17	.73
		1.32	.93	.17	.80
	h-l	-1.07	1.03	-.27	1.14
		-1.07	.62	.07	1.52
		.43	.96	.23	.57
		-.40	1.00	.07	.43
		.77	.63	.73	1.35
		1.43	.94	.40	.75

fatigue		Repetitiveness			
		non repetitive		repetitive	
		M	SD	M	SD
Sequence of DD	l-h	-1.03	1.25	-.17	.44
		-1.03	1.24	.00	.43
		.47	.77	.17	.27
		-.03	.42	.17	.48
		.30	.86	.00	.52
		.97	.67	.17	.63
	h-l	-.93	.71	-.22	.42
		-.77	.64	.45	.71
		-.60	.78	.12	.47
		-.10	.43	-.38	.55
		.40	.60	-.05	.39
		.57	.67	.45	.66

irritation		Repetitiveness			
		non repetitive		repetitive	
		M	SD	M	SD
Sequence of DD	l-h	-.20	.45	.05	.28
		-.20	.35	.05	.28
		.13	.44	.05	.66
		-.03	.50	.05	.40
		.13	.56	.22	.52
		.30	.68	.05	.45
	h-l	-.20	.26	.00	.00
		-.03	.48	.00	.00
		-.20	.26	.17	.41
		-.20	.43	.00	.00
		.13	.31	.17	.41
		.47	.46	.00	.00

strain		Repetitiveness			
		non repetitive		repetitive	
		M	SD	M	SD
Sequence of DD	l-h	.40	1.02	.12	.39
		.40	.85	-.05	.08
		.07	1.05	.12	.34
		-.27	.91	.12	.34
		.40	.49	-.05	.08
		.90	1.09	-.05	.08
	h-l	.43	.56	.00	.46
		.10	.70	.00	.46
		-.23	.33	.33	.34
		.10	.49	-.17	.37
		.43	.52	.33	.94
		-.07	.30	.17	.55

motivation		Repetitiveness			
		non repetitive		repetitive	
		M	SD	M	SD
Sequence of DD	l-h	1.13	.83	.15	.61
		.47	.78	-.52	.34
		-.37	.77	-.35	.41
		.13	.64	.32	.56
		-.20	.47	-.02	.60
		-1.03	.66	-.18	.53
	h-l	.62	.66	.80	.39
		.62	.66	.63	.45
		.45	.87	.13	.47
		.12	.33	-.03	.72
		-.38	.62	-.70	.93
		-.72	.48	-1.37	2.04

concentration		Repetitiveness			
		non repetitive		repetitive	
		M	SD	M	SD
Sequence of DD	l-h	1.42	.85	-.03	.63
		.42	.94	-.37	.33
		-.58	.69	-.20	.51
		.08	.53	.47	.46
		.08	.53	-.37	1.71
		-.75	.37	-.03	.74
	h-l	1.03	.53	.68	.26
		.37	.60	.35	.49
		-.13	.79	.18	.59
		.53	.73	-.15	.49
		-.30	.64	-.65	.94
		-.80	.33	-.32	.59

sleepiness		Repetitiveness			
		non repetitive		repetitive	
		M	SD	M	SD
Sequence of DD	l-h	-1.67	.73	.10	.77
		-.33	1.25	.27	1.27
		1.17	.84	-.07	.54
		-.33	.73	-.40	.81
		.50	1.06	-.40	1.05
		-.50	1.81	1.77	2.05
	h-l	-.80	.55	-.83	.67
		-.97	.59	-.50	.62
		-.13	.92	-.33	.66
		.03	.71	.17	.64
		.53	.41	1.17	1.46
		1.20	.97	2.17	1.72

Table B-35: Descriptive Statistics for SOF

			Repetitiveness			
			non repetitive		repetitive	
			M	SD	M	SD
Sequence of DD	l-h	Stress Run 1	1.90	.44	1.63	.28
		Stress Run 2	2.10	.51	1.70	.40
		Satiation Run 1	1.68	.30	1.93	.58
		Satiation Run 2	2.19	.57	2.00	.60
		Fatigue Run 1	2.22	.70	1.53	.25
		Fatigue Run 2	2.53	.48	1.77	.52
		Monotony Run 1	2.33	.33	2.48	.46
		Monotony Run 2	2.58	.45	2.45	.26
	h-l	Stress Run 1	1.92	.44	1.67	.23
		Stress Run 2	1.94	.42	1.75	.21
		Satiation Run 1	1.41	.35	1.57	.36
		Satiation Run 2	2.00	.66	2.07	.56
		Fatigue Run 1	1.92	.44	1.72	.34
		Fatigue Run 2	2.17	.31	1.90	.48
		Monotony Run 1	2.38	.38	2.22	.53
		Monotony Run 2	2.83	.40	2.75	.51

Table B-36: Descriptive Statistics for Nasa-TLX

			Repetitiveness			
			non repetitive		repetitive	
			M	SD	M	SD
Sequence of DD	l-h	mental demand_1	58	24	22	15
		mental demand_2	62	16	37	26
		physical demand_1	12	10	7	10
		physical demand_2	15	10	16	30
		temporal demand_1	47	23	13	16
		temporal demand_2	53	21	24	24
		performance_1	50	22	12	14
		performance_2	51	21	26	14
		effort_1	26	25	8	7
		effort_2	38	30	14	13
		frustration_1	10	5	12	18
		frustration_2	25	15	27	32
		feeling of monotony_1	38	26	84	13
		feeling of monotony_2	60	31	47	22
		situation awareness_1	67	26	83	7
		situation awareness_2	64	31	82	15
	h-l	mental demand_1	57	10	28	14
		mental demand_2	41	15	22	11
		physical demand_1	8	4	7	7
		physical demand_2	8	3	12	12
		temporal demand_1	20	25	22	15
		temporal demand_2	15	7	13	8
		performance_1	39	19	20	12
		performance_2	33	19	17	10
		effort_1	29	22	17	11
		effort_2	23	23	12	8
		frustration_1	15	11	7	9

	Repetitiveness			
	non repetitive		repetitive	
	M	SD	M	SD
frustration_2	18	13	12	4
feeling of monotony_1	44	20	53	13
feeling of monotony_2	57	22	69	20
situation awareness_1	72	15	81	12
situation awareness_2	74	20	78	12

B.2.4. ADDITIONAL RESULTS: PERFORMANCE DATA

Fisher exact Probability test

Two separate analyses were undertaken for the aggregated number of STCA alerts over two runs depending on repetititveness and dynamic density to compare the occurrence of critical states. The Fisher Exact Probability test was used because of the low sample size and the expected cell frequencies of not greater than 5. Calculations were executed with a program by Richard Lowry (1998), available at http://faculty.vassar.edu/lowry.

Table B-37: Descriptive Statistics and ANOVA results for VTS tests

a) Vienna Reaction Test

					Sequence of DD			
					l-h		h-l	
					M	SD	M	SD
Activity	non active	Repetitiveness	non repetitive	RT median reaction time	554	52	508	34
				RT median motor time	173	31	135	46
				RT distribution reaction time	23	7	16	2
				RT distribution motor time	18	8	22	9
			repetitive	RT median reaction time	574	70	549	57
				RT median motor time	169	47	149	27
				RT distribution reaction time	25	6	16	2
				RT distribution motor time	23	11	19	6
	active	Repetitiveness	non repetitive	RT median reaction time	493	74	550	108
				RT median motor time	105	11	158	33
				RT distribution reaction time	32	8	23	14
				RT distribution motor time	42	21	15	1
			repetitive	RT median reaction time	535	68	553	68
				RT median motor time	164	36	154	34
				RT distribution reaction time	21	3	22	8
				RT distribution motor time	24	15	16	4

Source	Dependent Variable	df	F	p
rep	RT median reaction time	1	.89	.360
	RT median motor time	1	1.34	.265
	RT distribution reaction time	1	.76	.398
	RT distribution motor time	1	.65	.433
DD	RT median reaction time	1	.00	.968
	RT median motor time	1	.08	.772
	RT distribution reaction time	1	4.21	.057
	RT distribution motor time	1	3.70	.072
activity	RT median reaction time	1	.22	.644
	RT median motor time	1	.63	.441
	RT distribution reaction time	1	2.36	.144
	RT distribution motor time	1	.74	.402
rep x DD	RT median reaction time	1	.03	.871
	RT median motor time	1	.66	.428
	RT distribution reaction time	1	.30	.588
	RT distribution motor time	1	.42	.525
rep x activity	RT median reaction time	1	.02	.885
	RT median motor time	1	.63	.441
	RT distribution reaction time	1	1.13	.303
	RT distribution motor time	1	1.16	.297
DD x activity	RT median reaction time	1	1.71	.209
	RT median motor time	1	3.22	.092
	RT distribution reaction time	1	.46	.509
	RT distribution motor time	1	3.61	.076
rep x DD x activity	RT median reaction time	1	.27	.607
	RT median motor time	1	2.07	.169
	RT distribution reaction time	1	1.03	.325
	RT distribution motor time	1	2.27	.151
Error	RT median reaction time	16		
	RT median motor time	16		
	RT distribution reaction time	16		
	RT distribution motor time	16		

b) Cognitrone (COG)

					Sequence of DD			
					l-h		h-l	
					M	SD	M	SD
Activity	non active	Repetitive-ness	non repetitive	Sum "correct reactions"	68	5	61	3
				Sum "incorrect reactions"	43	14	37	9
				Sum "incorrect non-reactions"	36	5	43	3
				M time "correct reactions" (sec)	1.26	.09	1.36	.05

			M time "incorrect reactions" (sec)	1.24	.09	1.35	.06
		repetitive	Sum "correct reactions"	64	12	66	0
			Sum "incorrect reactions"	36	14	39	13
			Sum "incorrect non-reactions"	40	12	38	0
			M time "correct reactions" (sec)	1.35	.10	1.33	.03
			M time "incorrect reactions" (sec)	1.31	.09	1.36	.07
active	Repetitive-ness	non repetitive	Sum "correct reactions"	57	13	70	10
			Sum "incorrect reactions"	28	4	39	8
			Sum "incorrect non-reactions"	47	13	34	10
			M time "correct reactions" (sec)	1.32	.07	1.30	.00
			M time "incorrect reactions" (sec)	1.27	.10	1.28	.01
		repetitive	Sum "correct reactions"	70	8	66	9
			Sum "incorrect reactions"	46	8	45	15
			Sum "incorrect non-reactions"	34	8	38	9
			M time "correct reactions" (sec)	1.28	.07	1.26	.04
			M time "incorrect reactions" (sec)	1.26	.08	1.29	.12

Source	Dependent Variable	df	F	p
rep	Sum "correct reactions"	1	.71	.411
	Sum "incorrect reactions"	1	1.11	.310
	Sum "incorrect non-reactions"	1	.71	.411
	M time "correct reactions" (sec)	1	.06	.809
	M time "incorrect reactions" (sec)	1	.47	.505
DD	Sum "correct reactions"	1	.07	.794
	Sum "incorrect reactions"	1	.12	.739
	Sum "incorrect non-reactions"	1	.07	.794
	M time "correct reactions" (sec)	1	.19	.672
	M time "incorrect reactions" (sec)	1	1.96	.182
activity	Sum "correct reactions"	1	.13	.722
	Sum "incorrect reactions"	1	.03	.871
	Sum "incorrect non-reactions"	1	.13	.722
	M time "correct reactions" (sec)	1	1.67	.215
	M time "incorrect reactions" (sec)	1	1.21	.289
rep x DD	Sum "correct reactions"	1	.36	.555
	Sum "incorrect reactions"	1	.02	.898
	Sum "incorrect non-reactions"	1	.36	.555
	M time "correct reactions"	1	1.37	.260

	(sec)			
	M time "incorrect reactions" (sec)	1	.06	.811
rep x activity	Sum "correct reactions"	1	.31	.587
	Sum "incorrect reactions"	1	2.26	.154
	Sum "incorrect non-reactions"	1	.31	.587
	M time "correct reactions" (sec)	1	1.67	.215
	M time "incorrect reactions" (sec)	1	.27	.610
DD x activity	Sum "correct reactions"	1	1.078	.316
	Sum "incorrect reactions"	1	.520	.482
	Sum "incorrect non-reactions"	1	1.08	.316
	M time "correct reactions" (sec)	1	1.23	.285
	M time "incorrect reactions" (sec)	1	.84	.375
rep x DD x activity	Sum "correct reactions"	1	3.46	.083
	Sum "incorrect reactions"	1	.99	.334
	Sum "incorrect non-reactions"	1	3.46	.083
	M time "correct reactions" (sec)	1	1.23	.285
	M time "incorrect reactions" (sec)	1	.35	.564
Error	Sum "correct reactions"	15		
	Sum "incorrect reactions"	15		
	Sum "incorrect non-reactions"	15		
	M time "correct reactions" (sec)	15		
	M time "incorrect reactions" (sec)	15		

c) ZBA

	Repetitiveness	Sequence of DD	Activity	M	SD	N
Median deviation time total ZBA	non repetitive	l-h	non active	1.57	1.00	3
			active	.60	.13	3
		h-l	non active	1.10	.59	3
			active	1.39	.35	3
	repetitive	l-h	non active	.77	.34	3
			active	1.04	.18	3
		h-l	non active	.55	.09	3
			active	.65	.31	3
Median direction deviation total ZBA	non repetitive	l-h	non active	63.67	21.78	3
			active	67.33	38.21	3
		h-l	non active	51.00	26.96	3
			active	51.00	9.64	3
	repetitive	l-h	non active	43.67	4.04	3
			active	55.00	19.31	3
		h-l	non active	52.33	28.00	3
			active	48.00	14.79	3

Source	Dependent Variable	F	df	p
rep	Median deviation time total ZBA	4.63	1	.047
	Median direction deviation total ZBA	.84	1	.374
DD	Median deviation time total ZBA	.14	1	.709
	Median direction deviation total ZBA	.54	1	.473
activity	Median deviation time total ZBA	.18	1	.678
	Median direction deviation total ZBA	.08	1	.778
rep x DD	Median deviation time total ZBA	1.51	1	.238
	Median direction deviation total ZBA	.68	1	.422
rep x activity	Median deviation time total ZBA	1.82	1	.196
	Median direction deviation total ZBA	.01	1	.930
DD x activity	Median deviation time total ZBA	2.04	1	.173
	Median direction deviation total ZBA	.28	1	.610
rep x DD x activity	Median deviation time total ZBA	3.47	1	.081
	Median direction deviation total ZBA	.10	1	.751
Error	Median deviation time total ZBA		16	
	Median direction deviation total ZBA		16	

B.2.5. CRITICAL STATES DISTINCTION

Table B-38 Analysis of variance for composed indicators of critical states

Source Between	Measure	F	df	p
rep	Monotony	12.514	1	.002
	Fatigue	.414	1	.527
	Satiation	2.624	1	.121
DD	Monotony	1.070	1	.313
	Fatigue	1.416	1	.248
	Satiation	.001	1	.971
rep x DD	Monotony	1.973	1	.175
	Fatigue	1.837	1	.190
	Satiation	.849	1	.368
Error	Monotony		20	
	Fatigue		20	
	Satiation		20	

Source Within	Measure	F	df	p
run	Monotony	.000	1	1.000
	Fatigue	.020	1	.890
	Satiation	.000	1	1.000
run x rep	Monotony	.100	1	.755
	Fatigue	4.284	1	.052
	Satiation	.003	1	.956
run x DD	Monotony	12.928	1	.002
	Fatigue	1.381	1	.254
	Satiation	2.925	1	.103
run x rep x DD	Monotony	3.660	1	.070
	Fatigue	.262	1	.615
	Satiation	1.017	1	.325
Error(run)	Monotony		20	
	Fatigue		20	
	Satiation		20	

B.2.6. Influence of additional variables

Table B-39: Descriptive statistics of personality indicators

a) Agreeableness

				Sequence of DD			
				l-h		h-l	
				M	SD	M	SD
HR (inv.)	low	non repetitive	Run 1	-8.23	1.78	-4.17	1.37
			Run 2	-6.57	1.56	-2.66	2.00
		repetitive	Run 1	-3.30	1.58	-3.89	2.12
			Run 2	.22	3.08	-1.03	.82
	high	non repetitive	Run 1	-2.79	3.11	-4.78	.
			Run 2	-4.03	.69	-3.72	.
		repetitive	Run 1	-4.15	1.04	-4.93	4.76
			Run 2	-2.97	1.24	-1.79	3.32
Sleepiness	low	non repetitive	Run 1	-.35	.36	-.73	.69
			Run 2	-.10	.45	.61	.36
		repetitive	Run 1	.92	.07	-.42	.40

				Sequence of DD			
				l-h		h-l	
				M	SD	M	SD
			Run 2	-.42	.07	.58	.54
	high	non repetitive	Run 1	-.13	.52	-.17	.
			Run 2	-.13	1.37	.50	.
		repetitive	Run 1	-.31	.56	-.63	.45
			Run 2	.69	.79	1.46	.33
Feeling of monotony	low	non repetitive	Run 1	29	25	45	22
			Run 2	49	33	62	20
		repetitive	Run 1	94	8	42	11
			Run 2	58	25	66	1
	high	non repetitive	Run 1	56	25	42	.
			Run 2	81	4	32	.
		repetitive	Run 1	79	12	59	11
			Run 2	41	22	70	26

b) Emotional Stability

				Sequence of DD			
				l-h		h-l	
				M	SD	M	SD
HR (inv.)	low	non repetitive	Run 1	-7.58	2.11	-3.89	.94
			Run 2	-6.16	1.63	-2.39	1.65
		repetitive	Run 1	-3.09	1.17	-5.24	3.91
			Run 2	-.74	2.74	-1.82	2.83
	high	non repetitive	Run 1	-.59	.	-6.17	.
			Run 2	-3.55	.	-5.07	.
		repetitive	Run 1	-4.64	.45	-1.32	.
			Run 2	-3.08	1.50	-.14	.
Sleepiness	low	non repetitive	Run 1	-.38	.32	-.73	.69
			Run 2	.09	.57	.54	.34
		repetitive	Run 1	.57	.61	-.53	.45
			Run 2	-.21	.36	1.07	.59
	high	non repetitive	Run 1	.23	.	-.17	.
			Run 2	-1.10	.	.83	.
		repetitive	Run 1	-.37	.67	-.70	.
			Run 2	.86	.88	1.63	.
Feeling of monotony	low	non repetitive	Run 1	38	30	43	22
			Run 2	56	33	58	24
		repetitive	Run 1	93	6	54	14
			Run 2	51	21	76	11
	high	non repetitive	Run 1	38	.	50	.
			Run 2	78	.	55	.
		repetitive	Run 1	74	10	50	.
			Run 2	42	27	33	.

c) Extraversion

				Sequence of DD			
				l-h		h-l	
				M	SD	M	SD
HR (inv.)	low	non repetitive	Run 1	-6.35	4.23	-3.41	1.44
			Run 2	-6.10	2.17	-2.94	1.13
		repetitive	Run 1	-3.65	1.27	-6.10	4.11
			Run 2	-1.22	3.47	-2.48	2.58
	high	non repetitive	Run 1	-6.54	2.19	-4.70	1.09
			Run 2	-4.96	.62	-2.79	2.28
		repetitive	Run 1	-4.09	1.27	-3.07	3.63
			Run 2	-2.60	.62	-.60	2.82
Sleepiness	low	non repetitive	Run 1	-.22	.47	-.80	.05
			Run 2	-.22	.65	.53	.52
		repetitive	Run 1	.19	.90	-.49	.31
			Run 2	.63	1.13	.96	.75
	high	non repetitive	Run 1	-.40	.14	-.55	.83
			Run 2	.10	1.04	.62	.29
		repetitive	Run 1	.01	.79	-.62	.55
			Run 2	.01	.42	1.38	.36
Feeling of monotony	low	non repetitive	Run 1	22	13	54	27
			Run 2	52	36	81	16
		repetitive	Run 1	83	5	45	9
			Run 2	62	19	74	14
	high	non repetitive	Run 1	69	6	40	18
			Run 2	74	13	46	12
		repetitive	Run 1	85	19	62	11
			Run 2	31	13	64	27

d) Conscientiousness

				Sequence of DD			
				l-h		h-l	
				M	SD	M	SD
HR (inv.)	low	non repetitive	Run 1	-5.98	3.90	-4.15	1.58
			Run 2	-6.18	2.13	-3.40	1.30
		repetitive	Run 1	-3.34	1.59	-4.82	3.43
			Run 2	-1.15	3.08	-2.03	2.23
	high	non repetitive	Run 1	-7.29	3.25	-4.51	.38
			Run 2	-4.80	.40	-1.71	2.84
		repetitive	Run 1	-4.40	.06	-4.47	4.54
			Run 2	-2.67	1.62	-1.29	3.10
Sleepiness	low	non repetitive	Run 1	-.33	.41	-.46	.40
			Run 2	-.49	.48	.54	.38
		repetitive	Run 1	.04	.85	-.08	.07
			Run 2	.04	.36	.58	.54
	high	non repetitive	Run 1	-.18	.45	-.98	1.15
			Run 2	.65	.26	.68	.26

				Sequence of DD			
				l-h		h-l	
				M	SD	M	SD
		repetitive	Run 1	.16	.86	-.79	.23
			Run 2	.60	1.18	1.46	.33
Feeling of monotony	low	non repetitive	Run 1	36	21	53	16
			Run 2	67	21	64	23
		repetitive	Run 1	81	16	61	16
			Run 2	43	29	75	14
	high	non repetitive	Run 1	41	46	28	20
			Run 2	46	53	45	18
		repetitive	Run 1	87	12	49	12
			Run 2	50	17	66	24

e) Intellect

				Sequence of DD			
				l-h		h-l	
				M	SD	M	SD
HR (inv.)	low	non repetitive	Run 1	-5.98	3.90	-3.00	.87
			Run 2	-6.18	2.13	-2.40	.36
		repetitive	Run 1	-3.59	1.39	-6.72	4.09
			Run 2	-1.99	3.03	-3.15	2.50
	high	non repetitive	Run 1	-7.29	3.25	-4.90	.87
			Run 2	-4.80	.40	-3.06	2.32
		repetitive	Run 1	-4.43	.02	-2.45	2.57
			Run 2	-1.74	.30	.07	1.80
Sleepiness	low	non repetitive	Run 1	-.33	.41	-.45	.54
			Run 2	-.49	.48	.22	.07
		repetitive	Run 1	.13	.71	-.27	.32
			Run 2	.13	.34	.96	.75
	high	non repetitive	Run 1	-.18	.45	-.73	.77
			Run 2	.65	.26	.78	.19
		repetitive	Run 1	.03	1.18	-.84	.25
			Run 2	.70	1.65	1.38	.36
Feeling of monotony	low	non repetitive	Run 1	36	21	63	15
			Run 2	67	21	65	38
		repetitive	Run 1	81	13	57	13
			Run 2	50	28	80	13
	high	non repetitive	Run 1	41	46	35	15
			Run 2	46	53	54	16
		repetitive	Run 1	90	14	49	15
			Run 2	40	0	58	22

Table B-40: ANOVA results of personality indicators

a) Agreeableness

Source		Wilks' Lambda	F	p	η_p^2
Between Subjects	rep	.444	5.85	.008	.556
	DD	.860	.76	.534	.140
	agreeableness	.908	.47	.707	.092
	rep * DD	.864	.74	.547	.136
	rep * agreeableness	.815	1.06	.397	.185
	DD * agreeableness	.906	.48	.699	.094
	rep * DD * agreeableness	.601	3.10	.061	.399
Within Subjects	run	.499	4.68	.018	.501
	run * rep	.683	2.17	.137	.317
	run * DD	.611	2.97	.068	.389
	run * agreeableness	.620	2.86	.075	.380
	run * rep * DD	.607	3.02	.065	.393
	run * rep * agreeableness	.695	2.05	.153	.305
	run * DD * agreeableness	.761	1.47	.267	.239
	run * rep * DD * agreeableness	.959	.19	.897	.041

Note. df hypothesis=3, df error=14.

b) Intellect

Source		Wilks' Lambda	F	p	η_p^2
Between Subjects	rep	.432	6.13	.007	.568
	DD	.837	.91	.460	.163
	intellect	.576	3.44	.046	.424
	rep * DD	.861	.75	.540	.139
	rep * intellect	.845	.85	.488	.155
	DD * intellect	.788	1.26	.327	.212
	rep * DD * intellect	.814	1.06	.396	.186
Within Subjects	run	.342	8.96	.001	.658
	run * rep	.601	3.10	.061	.399
	run * DD	.632	2.71	.085	.368
	run * intellect	.702	1.98	.163	.298
	run * rep * DD	.607	3.02	.065	.393
	run * rep * intellect	.904	.49	.693	.096
	run * DD * intellect	.876	.66	.590	.124
	run * rep * DD * intellect	.934	.33	.803	.066

Note. df hypothesis=3, df error=14.

c) Extraversion (extra)

Source		Wilks' Lambda	F	p	η_p^2
Between Subjects	rep	.395	7.14	.004	.605
	DD	.795	1.20	.345	.205
	extra1	.990	.05	.987	.010
	rep * DD	.857	.78	.526	.143
	rep * extra	.894	.55	.656	.106
	DD * extra	.736	1.68	.218	.264
	rep * DD * extra	.555	3.74	.037	.445
Within Subjects	run	.395	7.15	.004	.605
	run * rep	.566	3.57	.042	.434
	run * DD	.530	4.14	.027	.470
	run * extra	.609	2.99	.067	.391
	run * rep * DD	.594	3.18	.057	.406
	run * rep * extra	.874	.67	.584	.126
	run * DD * extra	.994	.03	.994	.006
	run * rep * DD * extra	.953	.23	.874	.047

Note. df hypothesis=3, df error=14.

d) Emotional stability (emot.)

Source		Wilks' Lambda	F	p	η_p^2
Between Subjects	rep	.567	3.57	.042	.433
	DD	.758	1.49	.261	.242
	emot1	.889	.58	.638	.111
	rep * DD	.939	.31	.821	.061
	rep * emot	.903	.504	.686	.097
	DD * emot	.874	.67	.584	.126
	rep * DD * emot	.645	2.57	.096	.355
Within Subjects	run	.555	3.75	.036	.445
	run * rep	.515	4.39	.022	.485
	run * DD	.565	3.59	.041	.435
	run * emot	.771	1.38	.289	.229
	run * rep * DD	.719	1.82	.190	.281
	run * rep * emot	.667	2.33	.118	.333
	run * DD * emot	.837	.91	.462	.163
	run * rep * DD * emot	.820	1.02	.412	.180

Note. df hypothesis=3, df error=14.

e) Conscientiousness (cons.)

Source		Wilks' Lambda	F	p	η_p^2
Between Subjects	rep	.419	6.46	.006	.581
	DD	.866	.72	.554	.134
	conscientiousness	.700	2.00	.160	.300
	rep * DD	.891	.57	.642	.109
	rep * conscientiousness	.940	.29	.827	.060
	DD * conscientiousness	.642	2.60	.093	.358
	rep * DD * conscientiousness	.843	.87	.482	.157
Within Subjects	run	.324	9.73	.001	.676
	run * rep	.677	2.22	.131	.323
	run * DD	.597	3.14	.059	.403
	run * conscientiousness	.717	1.84	.186	.283

Source	Wilks' Lambda	F	p	η_p^2
run * rep * DD	.691	2.09	.148	.309
run * rep * conscientiousness	.878	.65	.597	.122
run * DD * conscientiousness	.960	.19	.898	.040
run * rep * DD * conscientiousness	.914	.44	.728	.086

Note. df hypothesis=3, df error=14.

Table B-41 Descriptive statistics of ACQ, experience and MES

a) Experience

				Sequence of DD			
				l-h		h-l	
				M	SD	M	SD
HR (inv.)	low	non repetitive	Run 1	-5.78	3.93	-4.89	1.80
			Run 2	-5.26	1.72	-3.86	1.71
		repetitive	Run 1	-3.59	1.39	-6.32	1.31
			Run 2	-1.99	3.03	-2.61	1.41
	high	non repetitive	Run 1	-7.68	2.71	-3.96	1.07
			Run 2	-6.64	2.21	-2.33	1.90
		repetitive	Run 1	-4.43	.02	-3.72	4.59
			Run 2	-1.74	.30	-1.00	3.11
Sleepiness	low	non repetitive	Run 1	-.28	.35	-.12	.07
			Run 2	-.28	.83	.55	.40
		repetitive	Run 1	.13	.71	-.37	.47
			Run 2	.13	.34	.97	.00
	high	non repetitive	Run 1	-.28	.59	-.89	.68
			Run 2	.22	.35	.61	.35
		repetitive	Run 1	.03	1.18	-.65	.41
			Run 2	.70	1.65	1.27	.71
Feeling of monotony	low	non repetitive	Run 1	51	22	51	1
			Run 2	66	20	47	12
		repetitive	Run 1	81	13	53	27
			Run 2	50	28	76	13
	high	non repetitive	Run 1	12	5	41	24
			Run 2	47	54	63	25
		repetitive	Run 1	90	14	53	7
			Run 2	40	0	65	24

b) Morningness-Eveningness-Preference (mes)

				Sequence of DD			
				l-h		h-l	
				M	SD	M	SD
HR (inv.)	low	non repetitive	Run 1	-5.01	4.43	-4.61	1.09
			Run 2	-5.22	2.11	-2.79	2.28
		repetitive	Run 1	-3.89	1.54	-4.82	3.43

				Sequence of DD			
				l-h		h-l	
				M	SD	M	SD
			Run 2	-1.77	3.67	-2.03	2.23
	high	non repetitive	Run 1	-7.82	1.93	-3.58	1.69
			Run 2	-6.23	1.72	-2.93	1.11
		repetitive	Run 1	-3.85	1.01	-4.47	4.54
			Run 2	-2.05	.56	-1.29	3.10
Sleepiness	low	non repetitive	Run 1	-.27	.43	-.70	.80
			Run 2	-.16	.97	.72	.30
		repetitive	Run 1	.22	.85	-.08	.07
			Run 2	.11	.42	.58	.54
	high	non repetitive	Run 1	-.29	.42	-.50	.47
			Run 2	-.07	.55	.33	.24
		repetitive	Run 1	-.02	.84	-.79	.23
			Run 2	.53	1.20	1.46	.33
Feeling of monotony	low	non repetitive	Run 1	46	24	38	18
			Run 2	66	25	55	13
		repetitive	Run 1	77	13	61	16
			Run 2	54	32	75	14
	high	non repetitive	Run 1	29	31	58	22
			Run 2	53	40	62	42
		repetitive	Run 1	91	10	49	12
			Run 2	39	2	66	24

c) Performance-related action control style (AOP)

				Sequence of DD			
				l-h		h-l	
				M	SD	M	SD
HR (inv.)	low	non repetitive	Run 1	-8.23	1.78	-3.76	1.03
			Run 2	-6.57	1.56	-2.05	1.70
		repetitive	Run 1	-3.56	1.24	-3.95	4.67
			Run 2	-2.09	.79	-.82	3.94
	high	non repetitive	Run 1	-2.79	3.11	-5.30	1.23
			Run 2	-4.03	.69	-4.41	.94
		repetitive	Run 1	-4.02	1.28	-4.91	4.12
			Run 2	-1.82	3.00	-1.90	2.41
Sleepiness	low	non repetitive	Run 1	-.35	.36	-.72	.80
			Run 2	-.10	.45	.45	.31
		repetitive	Run 1	-.47	.47	-.58	.78
			Run 2	1.03	1.18	1.25	.40
	high	non repetitive	Run 1	-.13	.52	-.47	.42
			Run 2	-.13	1.37	.87	.05
		repetitive	Run 1	.38	.76	-.54	.27
			Run 2	-.03	.45	1.13	.70

				Sequence of DD			
				l-h		h-l	
				M	SD	M	SD
Feeling of monotony	low	non repetitive	Run 1	29	25	45	25
			Run 2	49	33	55	27
		repetitive	Run 1	86	8	68	6
			Run 2	39	2	79	8
	high	non repetitive	Run 1	56	25	43	11
			Run 2	81	4	63	11
		repetitive	Run 1	83	16	46	8
			Run 2	51	27	64	23

Table B-42: ANOVA results of ACQ, experience and MES

a) Experience (lic)

Source		Wilks' Lambda	F	p	η_p^2
Between Subjects	rep	.497	4.73	.018	.503
	DD	.855	.79	.519	.145
	lic	.917	.42	.742	.083
	rep * DD	.831	.94	.444	.169
	rep * lic	.924	.38	.767	.076
	DD * lic	.782	1.30	.312	.218
	rep * DD * lic	.878	.65	.597	.122
Within Subjects	run	.402	6.94	.004	.598
	run * rep	.619	2.87	.074	.381
	run * DD	.640	2.63	.091	.360
	run * lic	.897	.53	.667	.103
	run * rep * DD	.529	4.16	.027	.471
	run * rep * lic	.805	1.13	.371	.195
	run * DD * lic	.977	.11	.952	.023
	run * rep * DD * lic	.979	.09	.960	.021

b) Morningness-Eveningness-Preference (mes)

Source		Wilks' Lambda	F	p	η_p^2
Between Subjects	rep	.492	4.82	.017	.508
	DD	.843	.87	.481	.157
	mes	.986	.07	.976	.014
	rep * DD	.863	.74	.544	.137
	rep * mes	.957	.21	.889	.043
	DD * mes	.953	.23	.874	.047
	rep * DD * mes	.913	.44	.726	.087
Within Subjects	run	.416	6.54	.005	.584
	run * rep	.640	2.63	.091	.360
	run * DD	.626	2.79	.079	.374
	run * mes	.887	.59	.629	.113
	run * rep * DD	.603	3.07	.062	.397
	run * rep * mes	.847	.84	.494	.153
	run * DD * mes	.971	.14	.933	.029
	run * rep * DD * mes	.836	.91	.460	.164

c) Performance-related action control style (AOP)

Multivariate Tests(b)

Source		Wilks' Lambda	F	p	η_p^2
Between Subjects	rep	.439	5.97	.008	.561
	DD	.846	.85	.488	.154
	AOP	.943	.28	.839	.057
	rep * DD	.920	.40	.753	.080
	rep * AOP	.727	1.75	.202	.273
	DD * AOP	.710	1.90	.176	.290
	rep * DD * AOP	.894	.55	.655	.106
Within Subjects	run	.408	6.78	.005	.592
	run * rep	.569	3.53	.043	.431
	run * DD	.612	2.96	.069	.388
	run * AOP	.853	.80	.513	.147
	run * rep * DD	.651	2.50	.102	.349
	run * rep * AOP	.808	1.11	.379	.192
	run * DD * AOP	.915	.43	.734	.085
	run * rep * DD * AOP	.890	.58	.638	.110

B.2.7. RESULTS FOR DISCUSSION

B.2.7.1. Results for additional analysis

Table B-43: Additional descriptive statistics and ANOVA results for physiological indicators during performance tests

a) HR (corr. Baseline) during performance tests

				Repetitiveness					
				non repetitive			repetitive		
				Sequence of DD		Group Total	Sequence of DD		Group Total
				l-h	h-l		l-h	h-l	
Activity	non active	RT: HR corr	M	4.38	-1.64	1.37	-1.57	2.30	.36
			SD	6.33	4.45	5.90	5.23	2.57	4.25
		COG: HR corr	M	6.55	-.38	3.09	1.58	4.63	3.10
			SD	3.09	.45	4.28	3.13	1.96	2.87
		ZBA: HR corr	M	2.91	-2.53	.19	1.90	3.74	2.82
			SD	6.02	3.53	5.32	3.76	2.72	3.10
	active	RT: HR corr	M	7.55	4.36	5.96	-2.64	-1.05	-1.85
			SD	11.51	6.82	8.64	.99	1.11	1.28
		COG: HR corr	M	7.69	8.82	8.14	.91	3.35	2.13
			SD	4.91	.48	3.54	2.77	3.33	3.04
		ZBA: HR corr	M	5.60	3.14	4.37	.17	-.44	-.13
			SD	3.52	2.03	2.90	.93	1.46	1.15
Group Total	RT: HR corr	M		5.96	1.36	3.66	-2.10	.62	-.74
		SD		8.49	6.11	7.45	3.42	2.55	3.21
	COG: HR corr	M		7.12	3.30	5.38	1.25	3.99	2.62
		SD		3.72	5.05	4.60	2.67	2.54	2.87
	ZBA: HR corr	M		4.26	.31	2.28	1.04	1.65	1.34
		SD		4.65	4.04	4.64	2.62	3.01	2.71

Source	Dependent Variable	F	df	p
rep	RT: HR corr	4.712	1	.046
	COG: HR corr	5.953	1	.028
	ZBA: HR corr	.746	1	.401
DD	RT: HR corr	.003	1	.956
	COG: HR corr	.004	1	.950
	ZBA: HR corr	.940	1	.348
activity	RT: HR corr	.691	1	.419
	COG: HR corr	2.807	1	.115
	ZBA: HR corr	.405	1	.534
rep x DD	RT: HR corr	1.421	1	.252
	COG: HR corr	5.082	1	.040
	ZBA: HR corr	1.961	1	.182
rep x activity	RT: HR corr	3.071	1	.100

Source	Dependent Variable	F	df	p
	COG: HR corr	6.030	1	.027
	ZBA: HR corr	7.372	1	.016
DD x activity	RT: HR corr	.154	1	.701
	COG: HR corr	2.208	1	.158
	ZBA: HR corr	.089	1	.770
rep x DD x activity	RT: HR corr	.753	1	.399
	COG: HR corr	2.997	1	.104
	ZBA: HR corr	1.343	1	.265
Error	RT: HR corr		15	
	COG: HR corr		15	
	ZBA: HR corr		15	

b) HRV during performance tests

				Repetitiveness					
				non repetitive			repetitive		
				Sequence of DD		Group Total	Sequence of DD		Group Total
				l-h	h-l		l-h	h-l	
Activity	non active	RT: HRV corr	M	16.44	2.96	9.70	-38.17	17.64	-10.27
			SD	7.22	1.86	8.76	66.51	10.30	52.41
		COG: HRV corr	M	-.30	2.92	1.31	-53.82	-14.50	-34.16
			SD	7.90	4.03	5.88	76.52	5.63	53.09
		ZBA: HRV corr	M	7.10	1.59	4.35	-47.25	-5.04	-26.15
			SD	6.85	4.43	5.98	64.64	9.03	47.32
	active	RT: HRV corr	M	4.11	23.39	15.68	-15.59	-14.07	-14.83
			SD	.60	62.31	45.31	24.29	20.48	20.11
		COG: HRV corr	M	-2.05	-3.58	-2.81	-12.05	-16.88	-14.47
			SD	1.41	5.66	3.48	17.07	26.90	20.32
		ZBA: HRV corr	M	-2.10	-5.66	-4.24	-8.25	-18.87	-13.56
			SD	1.63	22.12	15.78	14.14	20.98	17.03
Group Total	RT: HRV corr	M		11.50	13.18	12.42	-26.88	1.78	-12.55
		SD		8.47	40.98	29.48	46.46	22.63	37.92
	COG: HRV corr	M		-1.00	.32	-.34	-32.94	-15.69	-24.31
		SD		5.71	5.37	5.27	54.61	17.43	39.68
	ZBA: HRV corr	M		3.42	-2.04	.44	-27.75	-11.96	-19.85
		SD		7.04	14.81	11.73	46.99	16.31	34.53

Source	Dependent Variable	F	df	p
rep	RT: HRV corr	4.760	1	.047
	COG: HRV corr	2.962	1	.107
	ZBA: HRV corr	4.072	1	.063
DD	RT: HRV corr	2.497	1	.136
	COG: HRV corr	.437	1	.519
	ZBA: HRV corr	.585	1	.457
activity	RT: HRV corr	.197	1	.664
	COG: HRV corr	.324	1	.579
	ZBA: HRV corr	.217	1	.649
rep x DD	RT: HRV corr	.203	1	.659
	COG: HRV corr	.359	1	.559
	ZBA: HRV corr	.365	1	.555
rep x activity	RT: HRV corr	.589	1	.456
	COG: HRV corr	.756	1	.399
	ZBA: HRV corr	.391	1	.542
DD x activity	RT: HRV corr	.006	1	.937
	COG: HRV corr	.797	1	.387
	ZBA: HRV corr	.682	1	.423
rep x DD x activity	RT: HRV corr	4.017	1	.065
	COG: HRV corr	.518	1	.484
	ZBA: HRV corr	2.143	1	.165
Error	RT: HRV corr		14	
	COG: HRV corr		14	
	ZBA: HRV corr		14	

Table B-44: Additional descriptive statistics and ANOVA results for rest break effects

			Repetitiveness					
			non repetitive			repetitive		
			Sequence of DD		Group Total	Sequence of DD		Group Total
			l-h	h-l		l-h	h-l	
Activity: non active	attentiveness (c.)	M	-.50	-.96	-.73	-.08	.25	.08
		SD	.38	1.25	.86	.44	.33	.39
	fatigue (c.)	M	1.67	2.00	1.83	.29	.13	.21
		SD	.52	1.52	1.03	.26	.54	.39
	boredom (c.)	M	.92	2.46	1.69	.04	-.25	-.10
		SD	.64	1.42	1.30	.29	1.21	.80
	irritation (c.)	M	.38	.38	.38	-.13	.00	-.06
		SD	.45	.22	.32	.13	.00	.10
	strain (c.)	M	-.50	-.04	-.27	.00	.13	.06
		SD	.70	.26	.53	.00	.13	.10
	concentration (c.)	M	-1.08	-1.21	-1.15	-.13	-.17	-.15
		SD	.31	1.15	.76	.25	.19	.20
	motivation (c.)	M	-1.46	-2.13	-1.79	-.21	-.42	-.31
		SD	.83	1.74	1.27	.29	.64	.46
	sleepiness (c.)	M	2.04	1.79	1.92	.33	-.13	.10
		SD	.26	1.63	1.05	.63	.76	.67

			Repetitiveness					
			non repetitive			repetitive		
			Sequence of DD		Group Total	Sequence of DD		Group Total
			l-h	h-l		l-h	h-l	
Activity: active	attentiveness (c.)	M	-.50	.00	-.25	.58	.08	.33
		SD	.98	.33	.71	.56	.26	.48
	fatigue (c.)	M	.33	.04	.19	-.08	-.46	-.27
		SD	.26	1.01	.68	.14	.40	.34
	boredom (c.)	M	.92	.08	.50	-.46	-.83	-.65
		SD	.19	1.39	1.00	1.04	.81	.86
	irritation (c.)	M	-.25	.17	-.04	-.17	-.42	-.29
		SD	.22	.52	.42	.29	.72	.51
	strain (c.)	M	-.38	-.33	-.35	-.13	-.54	-.33
		SD	.45	.94	.66	.13	.44	.37
	concentration (c.)	M	-.38	-.17	-.27	.25	.63	.44
		SD	.63	.88	.69	1.15	.25	.77
	motivation (c.)	M	-.92	-.08	-.50	-.25	.33	.04
		SD	.80	.26	.70	.88	.52	.72
	sleepiness (c.)	M	.25	-.04	.10	-.92	-.92	-.92
		SD	.66	.56	.57	.75	.56	.60

Source	Dependent Variable	F	df	p
rep	Attentiveness (c.)	6.76	1	.019
	fatigue (c.)	12.39	1	.003
	boredom (c.)	13.46	1	.002
	irritation (c.)	4.77	1	.044
	strain (c.)	.81	1	.382
	concentration (c.)	8.61	1	.010
	motivation (c.)	8.23	1	.011
	sleepiness (c.)	18.08	1	.001
DD	attentiveness (c.)	.01	1	.909
	fatigue (c.)	.18	1	.678
	boredom (c.)	.00	1	.980
	irritation (c.)	.21	1	.649
	strain (c.)	.07	1	.795
	concentration (c.)	.13	1	.725
	motivation (c.)	.15	1	.706
	sleepiness (c.)	.56	1	.464
activity	attentiveness (c.)	1.85	1	.193
	fatigue (c.)	12.89	1	.002
	boredom (c.)	4.66	1	.046
	irritation (c.)	4.22	1	.057
	strain (c.)	1.48	1	.241
	concentration (c.)	6.27	1	.023
	motivation (c.)	5.46	1	.033
	sleepiness (c.)	18.08	1	.001
rep x DD	attentiveness (c.)	.04	1	.849
	fatigue (c.)	.24	1	.629
	boredom (c.)	.74	1	.403
	irritation (c.)	.74	1	.402
	strain (c.)	1.01	1	.330
	concentration (c.)	.05	1	.833

Source	Dependent Variable	F	df	p
	motivation (c.)	.02	1	.884
	sleepiness (c.)	.00	1	.951
rep x activity	attentiveness (c.)	.18	1	.675
	fatigue (c.)	3.89	1	.066
	boredom (c.)	.65	1	.432
	irritation (c.)	.35	1	.559
	strain (c.)	.63	1	.439
	concentration (c.)	.25	1	.623
	motivation (c.)	1.77	1	.202
	sleepiness (c.)	1.41	1	.252
DD x activity	attentiveness (c.)	.01	1	.909
	fatigue (c.)	.49	1	.492
	boredom (c.)	2.36	1	.144
	irritation (c.)	.00	1	.948
	strain (c.)	1.48	1	.241
	concentration (c.)	.42	1	.529
	motivation (c.)	2.65	1	.123
	sleepiness (c.)	.09	1	.759
rep x DD x activity	attentiveness (c.)	2.78	1	.115
	fatigue (c.)	.12	1	.729
	boredom (c.)	2.05	1	.172
	irritation (c.)	1.58	1	.226
	strain (c.)	.03	1	.876
	concentration (c.)	.00	1	.944
	motivation (c.)	.25	1	.622
	sleepiness (c.)	.14	1	.712

B.2.7.2. Results for critical states analysis for hypothesis generating issues

Table B-45: Additional descriptive statistics and ANCOVA results for flow experience

Dependent Variable: Indicator of Monotony after Run 1(z-score)

Repetitiveness	Sequence of DD	M	SD	N
non repetitive	l-h	-.39	.67	6
	h-l	-.23	.57	6
	Total	-.31	.60	12
repetitive	l-h	.74	.63	6
	h-l	-.11	.50	6
	Total	.31	.70	12
Total	l-h	.17	.86	12
	h-l	-.17	.52	12
	Total	.00	.71	24

Correlation between Indicator of Monotony State and Flow

		Indicator of Monotony after Run 1 (z-score)	Flow indicator
Indicator of Monotony after Run 1 (z-score)	r	1	-.775(**)
	p		.000
	n	24	24
Flow indicator	r	-.775(**)	1
	p	.000	
	n	24	24

ANOCOVA: dependent Variable: Indicator of Monotony after Run 1(z-score), Covariate: flow

Source	F	df	p	η_p^2
flow	12.04	1	.003	.388
rep	.85	1	.369	.043
DD	.08	1	.778	.004
rep x DD	.86	1	.365	.043
Error		19		

Table B-46: Correlations (Pearson's r) between all indicators for each run

RUN 1		Workload	No. blinks	SCL	HR	HRV	Hedon. T.	Tense A.	Energetic A.	Sleepiness	Boredom	Fatigue	Motivation	Irritation	concentration	Monotony Ind.	SOF: monot.	SOF: fatigue	SOF: satiat.	SOF: stress
No. blinks	r	-.029	1	-.105	-.226	-.152	-.059	-.562	-.046	.170	.039	-.323	-.026	.025	.087	.175	.151	.346	.168	.285
	p	.895		.626	.288	.489	.784	.004	.830	.426	.855	.123	.904	.909	.685	.412	.481	.098	.433	.177
SCL	r	.039	-.105	1	.318	.352	-.107	.054	.549	-.181	.061	.160	-.267	.056	-.372	-.145	-.117	.091	.054	-.037
	p	.858	.626		.130	.099	.620	.803	.006	.397	.778	.456	.207	.795	.073	.499	.585	.672	.802	.862
HR	r	.408	-.226	.318	1	.148	-.033	.011	.501	-.024	-.099	.133	.041	.047	.171	-.628	-.190	.010	-.176	-.042
	p	.048	.288	.130		.500	.879	.959	.013	.910	.646	.537	.848	.828	.425	.001	.375	.962	.410	.847
HRV	r	-.050	-.152	.352	.148	1	.231	.219	.013	-.173	.150	.120	-.355	.108	-.318	-.040	.010	-.362	-.016	-.249
	p	.822	.489	.099	.500		.289	.316	.954	.429	.495	.585	.097	.625	.139	.856	.963	.090	.944	.252
Hedonic T.	r	.141	-.059	-.107	-.033	.231	1	-.084	-.468	.137	.330	.065	-.157	-.114	-.073	.115	.172	.089	.127	-.159
	p	.511	.784	.620	.879	.289		.698	.021	.523	.116	.763	.465	.594	.736	.592	.420	.679	.555	.458
Tense A.	r	-.311	-.562)	.054	.011	.219	-.084	1	-.102	-.076	.177	.242	-.137	.165	-.221	.137	.374	-.323	.254	-.149
	p	.139	.004	.803	.959	.316	.698		.635	.726	.407	.255	.522	.442	.299	.524	.072	.124	.232	.488
Energetic A.	r	.025	-.046	.549	.501	.013	-.468	-.102	1	-.239	-.116	.121	-.057	-.012	.078	-.445	-.333	.175	-.193	.098
	p	.907	.830	.006	.013	.954	.021	.635		.261	.589	.574	.792	.956	.718	.029	.112	.413	.366	.648
Sleepiness	r	-.168	.170	-.181	-.024	-.173	.137	-.076	-.239	1	.710	.434	-.526	-.021	-.548	.687	.389	.095	.534	.062
	p	.433	.426	.397	.910	.429	.523	.726	.261		.000	.034	.008	.921	.006	.000	.060	.658	.007	.772
Boredom	r	-.198	.039	.061	-.099	.150	.330	.177	-.116	.710	1	.644	-.658	-.009	-.561	.576	.391	-.031	.480	-.123
	p	.354	.855	.778	.646	.495	.116	.407	.589	.000		.001	.000	.966	.004	.003	.059	.885	.018	.567
Fatigue	r	-.044	-.323	.160	.133	.120	.065	.242	.121	.434	.644	1	-.401	-.008	-.474	.309	.166	-.218	.358	-.200

	p	.839	.123	.456	.537	.585	.763	.255	.574	.034	.001		.052	.969	.019	.142	.438	.307	.086	.349
Motivation	r	.142	-.026	-.267	.041	-.355	-.157	-.137	-.057	-.526	-.658	-.401	1	-.029	.610	-.474	-.194	-.119	-.323	-.197
	p	.507	.904	.207	.848	.097	.465	.522	.792	.008	.000	.052		.892	.002	.019	.364	.579	.124	.356
Irritation	r	-.139	.025	.056	.047	.108	-.114	.165	-.012	-.021	-.009	-.008	-.029	1	-.037	.109	-.079	-.246	-.111	-.156
	p	.516	.909	.795	.828	.625	.594	.442	.956	.921	.966	.969	.892		.862	.612	.712	.246	.606	.466
Concentrat.	r	.285	.087	-.372	.171	-.318	-.073	-.221	.078	-.548	-.561	-.474	.610)	-.037	1	-.652	-.309	.054	-.530	.039
	p	.176	.685	.073	.425	.139	.736	.299	.718	.006	.004	.019	.002	.862		.001	.141	.802	.008	.858
Monotony I.	r	-.552	.175	-.145	-.628	-.040	.115	.137	-.445	.687	.576	.309	-.474	.109	-.652	1	.507	-.044	.654	-.029
	p	.005	.412	.499	.001	.856	.592	.524	.029	.000	.003	.142	.019	.612	.001		.012	.840	.001	.891
SOF: monot.	r	-.101	.151	-.117	-.190	.010	.172	.374	-.333	.389	.391	.166	-.194	-.079	-.309	.507	1	.119	.742	.210
	p	.638	.481	.585	.375	.963	.420	.072	.112	.060	.059	.438	.364	.712	.141	.012		.579	.000	.325
SOF: fatigue	r	.175	.346	.091	.010	-.362	.089	-.323	.175	.095	-.031	-.218	-.119	-.246	.054	-.044	.119	1	.175	.842
	p	.414	.098	.672	.962	.090	.679	.124	.413	.658	.885	.307	.579	.246	.802	.840	.579		.414	.000
SOF: Satiat.	r	-.359	.168	.054	-.176	-.016	.127	.254	-.193	.534	.480	.358	-.323	-.111	-.530	.654	.742	.175	1	.109
	p	.085	.433	.802	.410	.944	.555	.232	.366	.007	.018	.086	.124	.606	.008	.001	.000	.414		.614
SOF: Stress	r	.312	.285	-.037	-.042	-.249	-.159	-.149	.098	.062	-.123	-.200	-.197	-.156	.039	-.029	.210	.842	.109	1
	p	.138	.177	.862	.847	.252	.458	.488	.648	.772	.567	.349	.356	.466	.858	.891	.325	.000	.614	

Run 2		Workload	No. blinks	SCL	HR	HRV	Hedon. T.	Tense A.	Energetic A.	Sleepiness	Boredom	Fatigue	Motivation	Irritation	concentration	Monotony Ind.	SOF: monot.	SOF: fatigue	SOF: satiat.	SOF: stress
No. blinks	r	-.175	1	-.217	.206	-.167	.333	-.455	-.001	-.302	.365	.335	-.184	.169	-.051	.022	.247	.471(x)	.286	.458(x)
	p	.412		.309	.335	.446	.111	.025	.997	.151	.080	.109	.391	.431	.812	.919	.244	.020	.176	.025
SCL	r	-.099	-.217	1	-.007	.239	.144	.339	.010	-.045	-.047	.167	.060	-.187	.187	.068	.225	-.121	.255	-.138
	p	.644	.309		.973	.272	.501	.106	.964	.836	.828	.436	.780	.382	.383	.751	.290	.574	.230	.519
HR	r	.463	.206	-.007	1	-.177	.191	-.186	.356	-.228	.279	.274	-.032	.114	-.185	-.660	-.129	.395	-.111	.295
	p	.023	.335	.973		.420	.371	.385	.088	.285	.186	.196	.882	.596	.386	.000	.549	.056	.606	.162
HRV	r	.322	-.167	.239	-.177	1	-.046	.055	-.094	.483	-.047	.071	.215	-.020	-.225	.356	.039	-.289	-.080	-.334
	p	.134	.446	.272	.420		.834	.803	.671	.020	.833	.749	.325	.929	.303	.096	.859	.181	.715	.119
Hedonic T.	r	.089	.333	.144	.191	-.046	1	.001	-.117	-.621	-.297	.375	.184	-.088	.079	-.428	-.018	.268	.236	.272
	p	.680	.111	.501	.371	.834		.995	.587	.001	.159	.071	.390	.681	.714	.037	.935	.205	.267	.198
Tense A.	r	-.157	-.455	.339	-.186	.055	.001	1	-.007	-.101	-.347	-.265	.278	-.061	.075	.084	.381	-.250	.228	-.349
	p	.464	.025	.106	.385	.803	.995		.974	.640	.096	.211	.189	.778	.728	.697	.066	.238	.285	.094
Energetic A.	r	.317	-.001	.010	.356	-.094	-.117	-.007	1	-.085	.170	-.087	.054	.052	.126	-.292	-.174	-.174	-.128	-.147
	p	.131	.997	.964	.088	.671	.587	.974		.692	.426	.686	.803	.808	.557	.166	.416	.417	.551	.492
Sleepiness	r	.113	-.302	-.045	-.228	.483	-.621	-.101	-.085	1	.229	-.086	-.350	-.022	-.256	.660	-.042	-.445	-.387	-.348
	p	.600	.151	.836	.285	.020	.001	.640	.692		.282	.688	.094	.920	.227	.000	.844	.029	.061	.096
Boredom	r	.030	.365	-.047	.279	-.047	-.297	-.347	.170	.229	1	.428	-.365	-.137	-.423	.172	.309	.458	.226	.389
	p	.888	.080	.828	.186	.833	.159	.096	.426	.282		.037	.080	.525	.039	.422	.142	.025	.287	.061
Fatigue	r	.217	.335	.167	.274	.071	.375	-.265	-.087	-.086	.428	1	-.035	-.177	-.013	-.147	.101	.380	.029	.563
	p	.308	.109	.436	.196	.749	.071	.211	.686	.688	.037		.873	.409	.951	.493	.638	.067	.894	.004
Motivation	r	.290	-.184	.060	-.032	.215	.184	.278	.054	-.350	-.365	-.035	1	-.179	-.002	-.286	-.221	.006	-.077	-.004
	p	.169	.391	.780	.882	.325	.390	.189	.803	.094	.080	.873		.402	.993	.176	.299	.976	.721	.984
Irritation	r	.114	.169	-.187	.114	-.020	-.088	-.061	.052	-.022	-.137	-.177	-.179	1	.102	-.053	.144	.146	.074	.052
	p	.597	.431	.382	.596	.929	.681	.778	.808	.920	.525	.409	.402		.634	.806	.501	.498	.732	.810
Concentrat.	r	-.211	-.051	.187	-.185	-.225	.079	.075	.126	-.256	-.423	-.013	-.002	.102	1	-.203	-.275	-.313	-.176	-.011
	p	.322	.812	.383	.386	.303	.714	.728	.557	.227	.039	.951	.993	.634		.340	.193	.136	.411	.961
Monotony	r	-.433	.022	.068	-.660	.356	-.428	.084	-.292	.660	.172	-.147	-.286	-.053	-.203	1	.428	-.385	.093	-.341

I.																				
	p	.034	.919	.751	.000	.096	.037	.697	.166	.000	.422	.493	.176	.806	.340		.037	.063	.664	.103
SOF: monot.	r	-.446	.247	.225	-.129	.039	-.018	.381	-.174	-.042	.309	.101	-.221	.144	-.275	.428	1	.253	.751	.047
	p	.029	.244	.290	.549	.859	.935	.066	.416	.844	.142	.638	.299	.501	.193	.037		.233	.000	.826
SOF: fatigue	r	.144	.471	-.121	.395	-.289	.268	-.250	-.174	-.445	.458	.380	.006	.146	-.313	-.385	.253	1	.442	.796
	p	.501	.020	.574	.056	.181	.205	.238	.417	.029	.025	.067	.976	.498	.136	.063	.233		.031	.000
SOF: Satiat.	r	-.339	.286	.255	-.111	-.080	.236	.228	-.128	-.387	.226	.029	-.077	.074	-.176	.093	.751	.442	1	.145
	p	.105	.176	.230	.606	.715	.267	.285	.551	.061	.287	.894	.721	.732	.411	.664	.000	.031		.499
SOF: Stress	r	.063	.458	-.138	.295	-.334	.272	-.349	-.147	-.348	.389	.563	-.004	.052	-.011	-.341	.047	.796	.145	1
	p	.769	.025	.519	.162	.119	.198	.094	.492	.096	.061	.004	.984	.810	.961	.103	.826	.000	.499	

Note. N=24; HRV: n=23

Table B-47
Pearson Correlations between values in critical state subscales (SOF) and the composed indicator of monotony for each run (for hypothesis generating purposes). To compare the course of all indicators that might contribute to understand the relationship between different critical states, the indicator for monotony was correlated with the scores in the SOF subscales. The results indicated that the composed monotony indicator significantly correlated with monotony and satiation in the first run, but only with monotony in the second run.

Subscale			Monotony I	Monotony	Fatigue	Satiation	Stress
Run 1	Monotony I 1	r	1	.507(*)	-.044	.654(**)	-.029
		p		.012	.840	.001	.891
	Monotony 1	r			.119	.742(**)	.210
		p			.579	.000	.325
	Fatigue 1	r				.175	.842(**)
		p				.414	.000
	Satiation 1	r					.109
		p					.614
	Stress 1	r					1
		p					
Run 2	Monotony I 2	r	1	.428(*)	-.385	.093	-.341
		p		.037	.063	.664	.103
	Monotony 2	r			.253	.751(**)	.047
		p			.233	.000	.826
	Fatigue 2	r				.442(*)	.796(**)
		p				.031	.000
	Satiation 2	r					.145
		p					.499
	Stress 2	r					1
		p					

Note. N=24 . ***p<.001. **p<.01. *p<.05. (*)p<0.1.

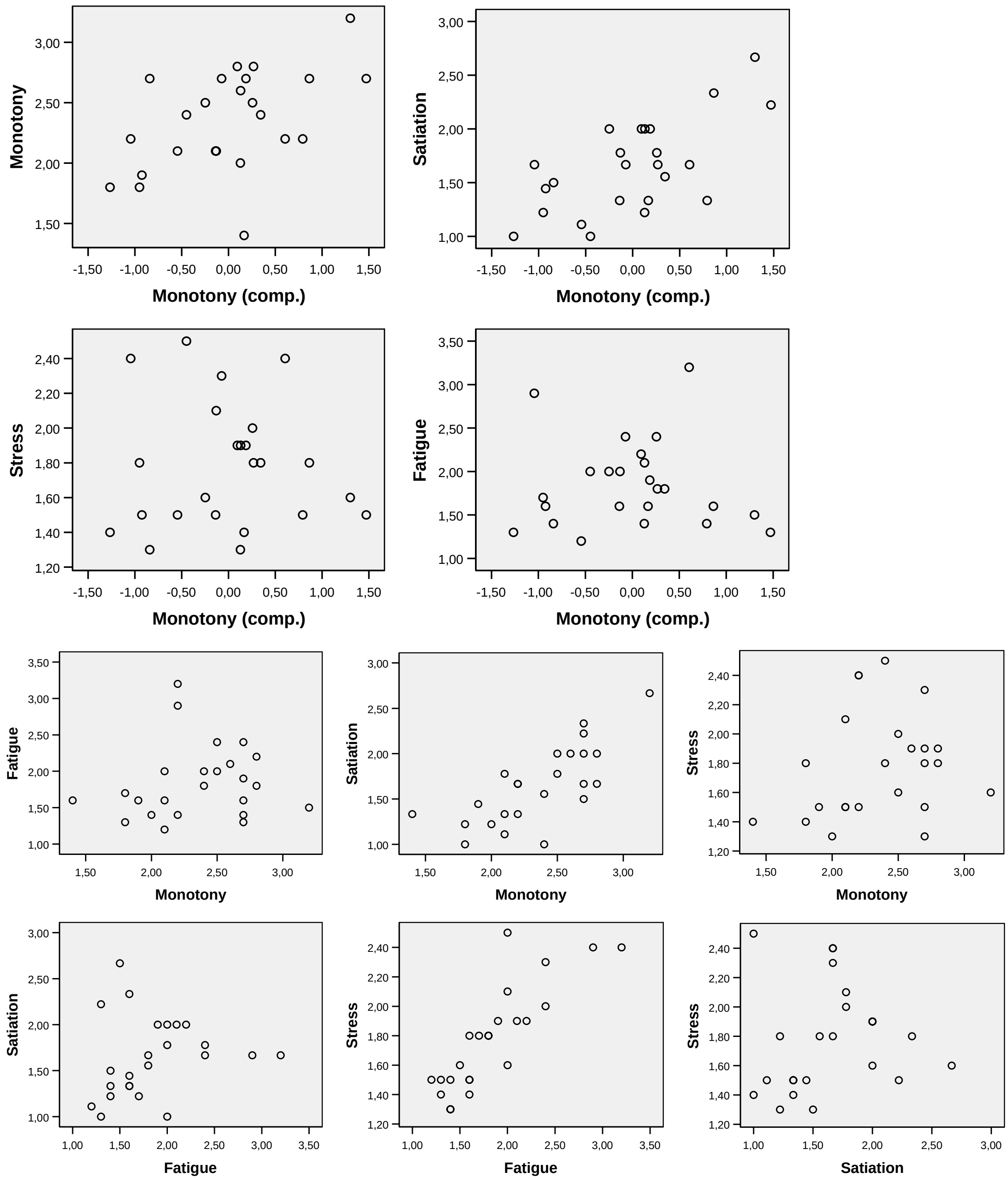


Figure B-1. Scatterplots between SOF criticalstate subsales and the composed indicator of monotony for Run1.

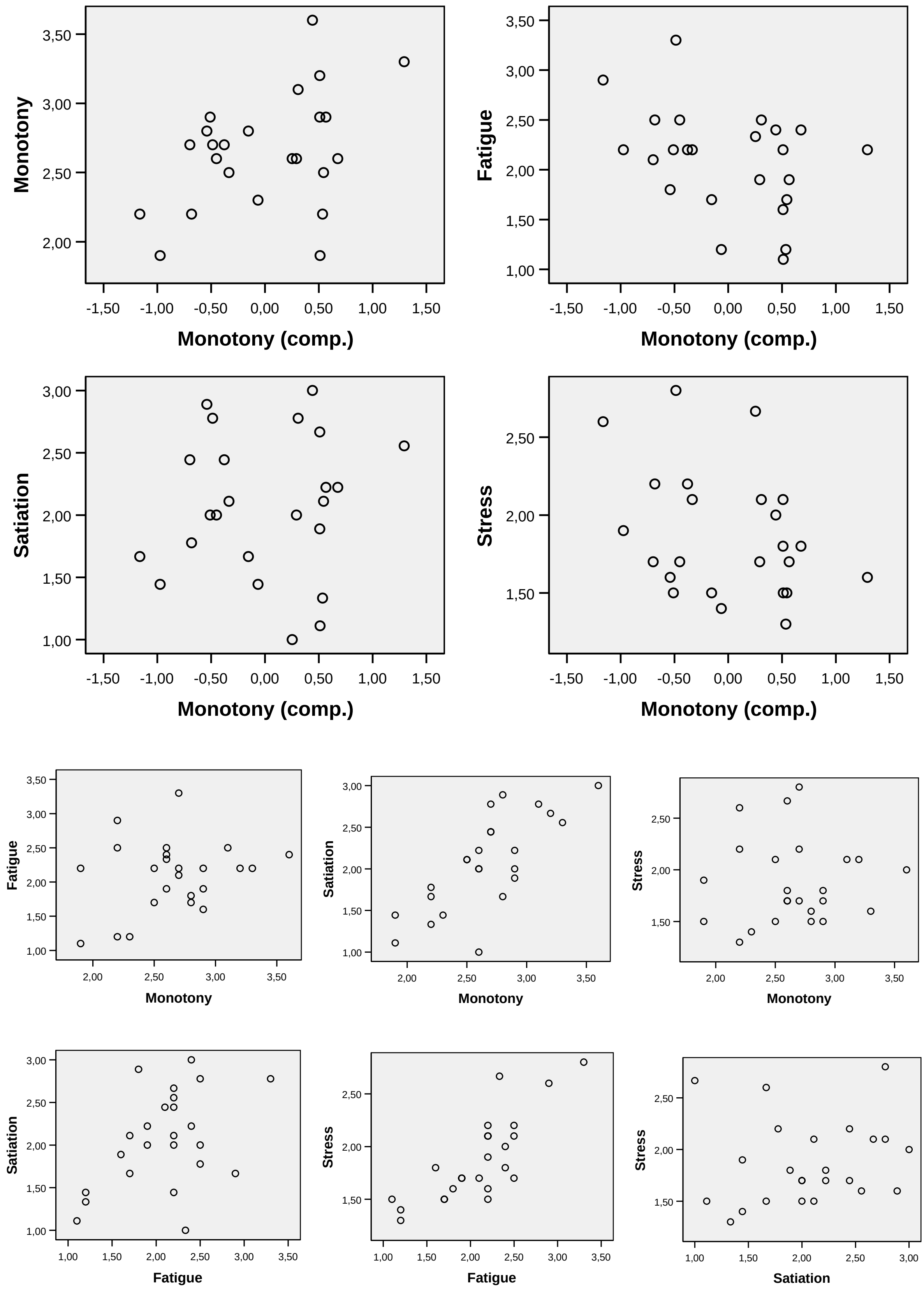


Figure B-2. Scatterplots between SOF criticalstate subsales and the composed indicator of monotony for Run2.

It was assumed that if concepts are independent, then there should be no correlations between the different indicators. As indicated in the correlations between these formed scales, there is no correlation between the formed scales. There is a correlation between the formed fatigue scale and the SOF fatigue scale; there is no correlation between the formed satiation items and the SOF satiation scale. The correlation is also present for the monotony indicator and the monotony subscale. There was no correlation between the satiation items and the SOF satiation scale.

To find additional support for the description of states, the relationships with other items were investigated. It was assumed that if monotony and satiation were independent, then boredom should increase and motivation decrease in satiation and not be connected with monotony. This was however not supported by the data (Table B-29). In the first run, motivation was significantly lower with increased scores in the fatigue and the monotony indicator correlated negatively with boredom. In the second run, fatigue correlated negatively with boredom and satiation correlated negatively with motivation.

Table B-48
Correlation between composed new indicators for critical states and ratings of boredom and motivation for each run (Study II). **To verify the procedure of composing the items based on Richter et al. (2002), correlations were calculated with the SOF subscales for hypothesis generating purposes.**

Critical State Indicators			Satiation	Fatigue	Monotony	Motivation	Boredom
Run 1	Satiation	*r*	1	-.232	.189	.194	-.055
		p		.275	.377	.363	.797
	Fatigue	r			.235	-.469(*)	.420(*)
		p			.270	.021	.041
	Monotony	*r*				-.474(*)	.576(**)
		p				.019	.003
	Motivation	*r*					-.658(**)
		p					.000
	Boredom	*r*					1
		p					
Run 2	Satiation	*r*		-.282	.165	-.415(*)	-.175
		p		.182	.442	.044	.415
	Fatigue	r			-.124	-.058	.682(**)
		p			.564	.788	.000
	Monotony	*r*				-.286	.172
		p				.176	.422
	Motivation	*r*					-.365
		p					.080
	Boredom	*r*					1
		p					

Note. N=24 . Pearsons Correlation. Composed Indicators standardized. ***p<.001. **p<.01. *p<.05. (*)p<0.1.

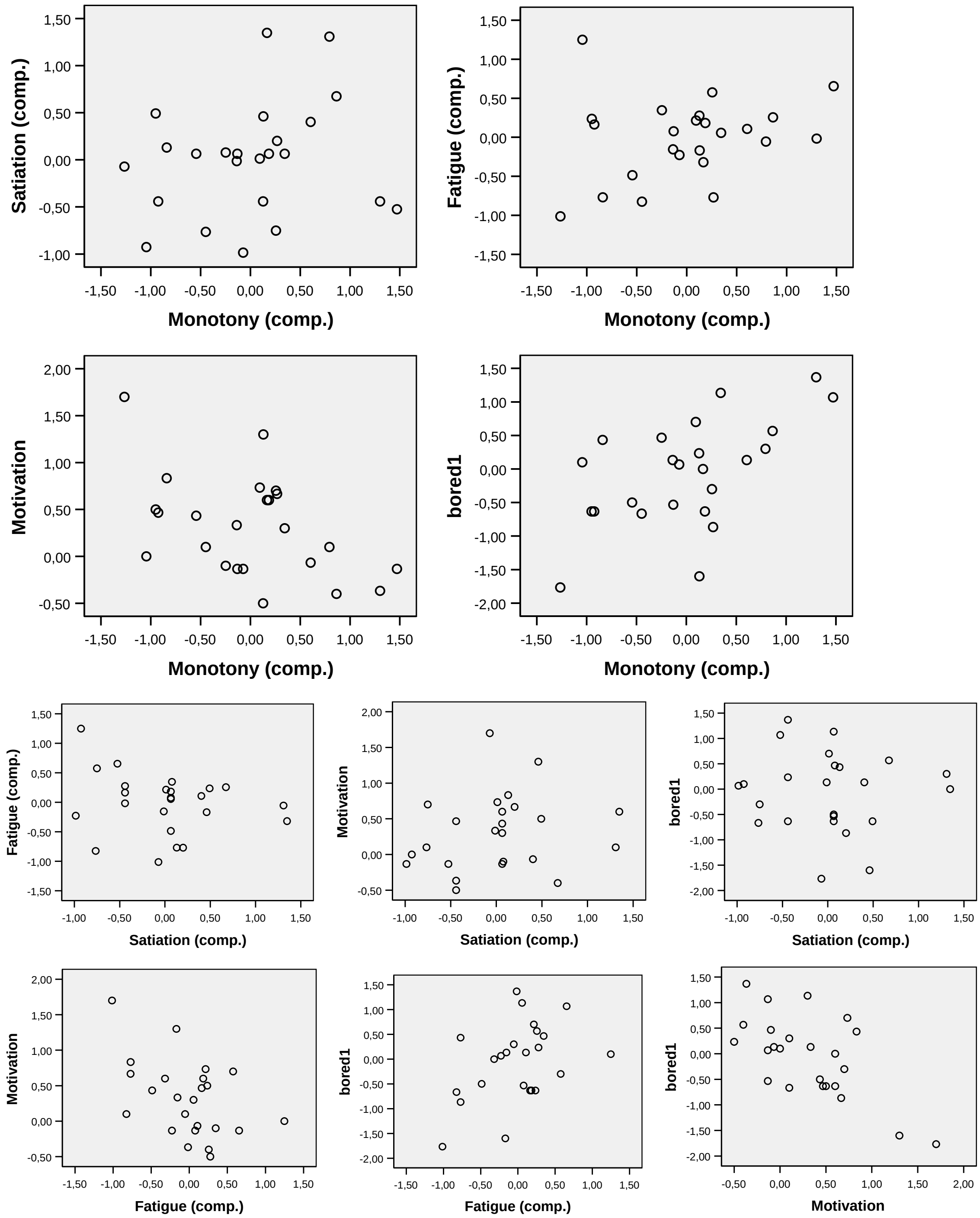


Figure B-3. Scatterplots between composed critical state indicators for monotony, fatigue and satiation and ratings for motivation and boredom in Run1.

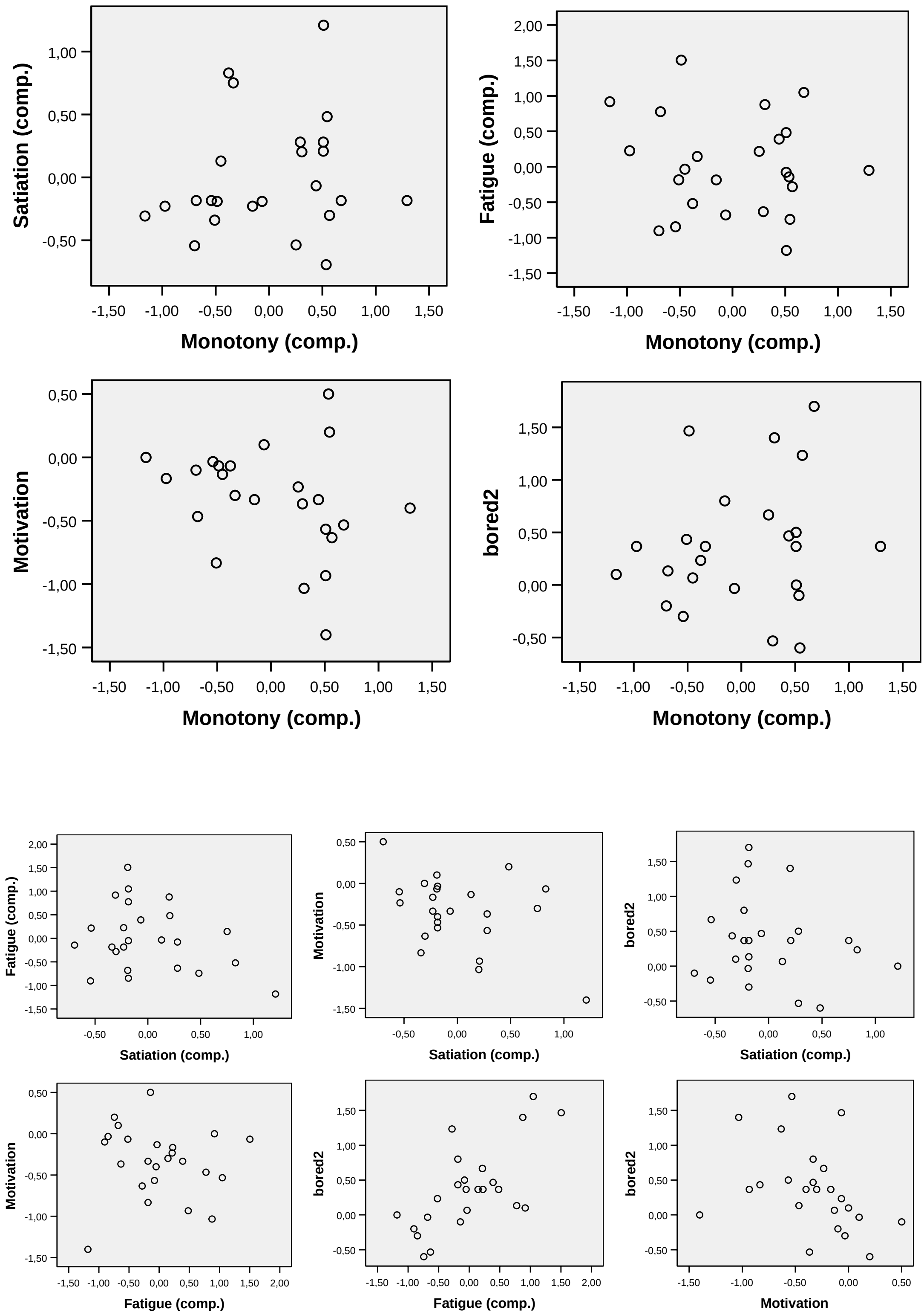


Figure B-4. Scatterplots between composed critical state indicators for monotony, fatigue and satiation and ratings for motivation and boredom in Run 2.

Appendix C: STUDY III

C.1. RESULTS CONFIRMATIVE HYPOTHESIS

Table C-1: Descriptive statistics for confirmative hypothesis 1

			Traffic Load				Group Total	
			low		high			
			M	SD	M	SD	M	SD
Repetitiveness	non repetitive	Feeling of Monotony (z-value)	-.40	1.08	-.47	.63	-.43	.86
		Sleepiness (z-value)	-.35	1.01	-.43	.72	-.39	.86
		HR inv. (z-value)	-.27	1.06	.15	1.12	-.04	1.08
		Monotony indicator (z-value)	-.37	.88	-.23	.69	-.30	.77
	repetitive	Feeling of Monotony (z-value)	.89	.90	-.02	.81	.43	.95
		Sleepiness (z-value)	.76	1.07	.02	.82	.39	1.00
		HR inv. (z-value)	.12	.92	-.02	1.01	.04	.95
		Monotony indicator (z-value)	.62	.66	-.01	.64	.31	.71
Group Total	Feeling of Monotony (z-value)		.24	1.17	-.24	.74	.00	1.00
	Sleepiness (z-value)		.20	1.16	-.20	.79	.00	1.00
	HR inv. (z-value)		-.07	.98	.06	1.04	.00	1.00
	Monotony indicator (z-value)		.12	.91	-.12	.66	.00	.79

Table C-2: Descriptive statistics for confirmative hypothesis 2

Monotony indicator (z-value)				Traffic Load				Group Total	
				low		high			
				M	SD	M	SD	M	SD
initial recover y	low	Repetitiveness	non repetitive	.06	.84	.10	.66	.08	.72
			repetitive	.80	.38	.30	.62	.55	.55
		Group Total		.43	.73	.20	.62	.32	.67
	high	Repetitiveness	non repetitive	-1.02	.44	-.74	.36	-.88	.40
			repetitive	.35	.95	-.47	.36	-.06	.79
		Group Total		-.33	1.00	-.60	.36	-.47	.74

Monotony indicator (z-value)			Traffic Load				Group Total	
			low		high			
			M	SD	M	SD	M	SD
low strain	Repetitiveness	non repetitive	-.62	.46	-.39	.64	-.50	.54
		repetitive	.51	.94	-.10	.76	.21	.87
	Group Total		-.05	.92	-.24	.68	-.15	.79
high strain	Repetitiveness	non repetitive	-.12	1.17	-.08	.78	-.10	.94
		repetitive	.72	.22	.09	.58	.40	.53
	Group Total		.30	.91	.00	.65	.15	.79

Monotony indicator (z-value)				Traffic Load					
				low		high		Group Total	
				M	SD	M	SD	M	SD
boredom proneness group	low	Repetitiveness	non repetitive	-.42	1.07	-.42	.71	-.42	.86
			repetitive	.40	.80	-.17	.53	.11	.71
		Group Total		-.01	.99	-.30	.60	-.15	.81
	high	Repetitiveness	non repetitive	-.25	.87	-.22	.65	-.23	.71
			repetitive	.70	.37	-.07	.66	.32	.64
		Group Total		.22	.80	-.14	.61	.04	.71

C.2. RESULTS DESCRIPTIVE HYPOTHESIS

Table C-3: Descriptive statistics for descriptive hypothesis 3: subjective TSI ratings

			Traffic Load				Group Total	
			low		high			
			M	SD	M	SD	M	SD
Repetitiveness	non repetitive	attentiveness	4.15	1.17	4.50	1.36	4.33	1.25
		sleepiness	2.33	1.79	2.20	1.28	2.27	1.52
		boredom	2.63	2.00	2.37	1.22	2.51	1.64
		strain	3.43	1.44	4.27	.93	3.85	1.25
		concentration	4.60	1.12	4.77	1.34	4.68	1.21
		motivation	4.07	1.64	4.73	1.27	4.40	1.47
		flow	2.25	1.86	2.15	1.11	2.20	1.49
		Confidence	6	1	5	1	5	1
	repetitive	attentiveness	2.83	1.19	3.53	.92	3.18	1.09
		sleepiness	4.30	1.89	3.00	1.46	3.65	1.77
		boredom	5.00	1.81	3.20	1.42	4.10	1.84
		strain	2.10	1.35	2.87	1.08	2.48	1.25
		concentration	2.57	1.22	3.57	.79	3.07	1.12
		motivation	3.07	1.40	4.27	1.39	3.67	1.49
		flow	1.25	1.89	2.20	1.55	1.73	1.75
		Confidence	5	1	6	1	5	1
Group Total	attentiveness		3.46	1.33	4.02	1.24	3.74	1.30
	sleepiness		3.32	2.06	2.60	1.40	2.96	1.77
	boredom		3.82	2.22	2.81	1.36	3.32	1.90
	strain		2.77	1.52	3.57	1.21	3.17	1.42
	concentration		3.58	1.54	4.17	1.24	3.88	1.41
	motivation		3.57	1.57	4.50	1.32	4.03	1.51
	flow		1.75	1.90	2.18	1.31	1.96	1.62
	Confidence		5.10	1.29	5.45	1.28	5.28	1.28

Table C-4: Descriptive statistics for descriptive hypothesis 3: Nasa-TLX ratings

			Traffic Load				Group Total	
			low		high			
			M	SD	M	SD	M	SD
Repetitiveness	non repetitive	mental demand	4.37	1.53	4.87	1.01	4.62	1.29
		temporal demand	3.63	1.36	4.20	1.43	3.92	1.39
		effort	4.10	1.61	4.63	1.06	4.37	1.35
		performance	5.53	.71	5.17	1.34	5.35	1.06
		frustration	2.27	1.84	2.37	1.41	2.32	1.60
		workload	3.43	.70	3.83	.56	3.63	.65
	repetitive	mental demand	2.47	1.23	3.80	1.47	3.13	1.48
		temporal demand	1.33	.72	2.53	1.15	1.93	1.12
		effort	1.93	1.29	3.00	1.02	2.47	1.26
		performance	4.50	1.54	4.97	1.32	4.73	1.42
		frustration	2.50	1.43	2.60	1.18	2.55	1.28
		workload	2.25	.69	2.98	.64	2.61	.75
Group Total	mental demand		3.42	1.66	4.33	1.34	3.88	1.56
	temporal demand		2.48	1.59	3.37	1.53	2.93	1.60
	effort		3.02	1.80	3.82	1.31	3.42	1.61
	performance		5.02	1.28	5.07	1.30	5.04	1.27
	frustration		2.38	1.61	2.48	1.27	2.43	1.43
	workload		2.84	.91	3.41	.73	3.12	.86

Table C-5: Descriptive statistics for descriptive hypothesis 3: SOF subscales and UWIST

			Traffic Load				Group Total	
			low		high		M	SD
			M	SD	M	SD		
Repetitiveness	non repetitive	SOF: Monotony	2.02	.32	2.06	.23	2.04	.27
		SOF: Fatigue	1.76	.50	1.83	.43	1.80	.45
		SOF: Stress	1.68	.51	1.59	.24	1.63	.39
		SOF: Satiation	1.73	.57	1.63	.47	1.68	.51
	repetitive	SOF: Monotony	2.19	.34	1.92	.24	2.05	.32
		SOF: Fatigue	1.81	.25	1.61	.49	1.71	.40
		SOF: Stress	1.51	.30	1.51	.39	1.51	.35
		SOF: Satiation	1.70	.42	1.51	.42	1.60	.42
Group Total	SOF: Monotony		2.10	.33	1.99	.24	2.04	.29
	SOF: Fatigue		1.78	.39	1.72	.46	1.75	.42
	SOF: Stress		1.60	.42	1.55	.32	1.57	.37
	SOF: Satiation		1.72	.49	1.57	.44	1.64	.47

			Traffic Load				Group Total	
			low		high		M	SD
			M	SD	M	SD		
Repetitiveness	non repetitive	SOF: Hedonism	2.55	.21	2.61	.25	2.58	.23
		SOF: Tense	2.83	.20	2.80	.20	2.82	.19
		SOF: Energy	2.11	.42	2.09	.29	2.10	.36
	repetitive	SOF: Hedonism	2.56	.28	2.56	.19	2.56	.23
		SOF: Tense	2.83	.18	2.81	.21	2.82	.19
		SOF: Energy	2.00	.21	2.01	.44	2.01	.34
Group Total	SOF: Hedonism		2.55	.24	2.59	.22	2.57	.23
	SOF: Tense		2.83	.18	2.80	.20	2.82	.19
	SOF: Energy		2.06	.33	2.05	.36	2.06	.34

Table C-6: Descriptive statistics for performance ratings

			Traffic Load				Group Total	
			low		high			
			M	SD	M	SD	M	SD
Rep	non repetitive	not feel focused anymore	.40	.70	.40	.52	.40	.60
		feel complacent	1.78	2.05	2.67	2.12	2.22	2.07
		ask to repeat call from a/c	.90	1.10	.70	.82	.80	.95
		not understand R/T	.50	.85	.22	.44	.37	.68
		spot a conflict only 1-6 minutes before	.50	.85	.60	.97	.55	.89
		feel like working less precise	.80	1.03	.33	.50	.58	.84
		make small mistakes (e.g. input errors)	.70	.82	.60	.52	.65	.67
		feel like getting behind in work	.10	.32	.20	.42	.15	.37
		feel like doing less pre-planning	.60	.70	.56	.73	.58	.69
		work slower as usual	.20	.42	.50	.85	.35	.67
		not knowing a/c on frequency	.30	.48	.30	.48	.30	.47
		forgetting routine co-ordination	.00	.00	.10	.32	.05	.22
		react slower as usual	.60	.52	.40	.70	.50	.61
		pay less attention to detail	.50	.71	.50	.53	.50	.61
		surprised by call	.20	.42	.40	.52	.30	.47
		miss a call	.30	.67	.20	.42	.25	.55
		feel easily distracted	.40	.52	.33	.50	.37	.50
		look for traffic that calls in	.80	.79	.80	.79	.80	.77
		overlooking obvious problems	.30	.48	.20	.42	.25	.44
		do less scanning	.40	.52	.50	.53	.45	.51
	repetitive	not feel focused anymore	.80	.79	.70	1.06	.75	.91
		feel complacent	1.80	1.93	2.70	2.06	2.25	2.00
		ask to repeat call from a/c	.40	.52	1.00	.67	.70	.66
		not understand R/T	.20	.42	.50	.53	.35	.49
		spot a conflict only 1-6 minutes before	.80	1.93	.50	1.08	.65	1.53
		feel like working less	.50	.71	1.20	1.81	.85	1.39

		Traffic Load				Group Total	
		low		high			
		M	SD	M	SD	M	SD
	precise						
	make small mistakes (e.g. input errors)	.40	.52	.50	.53	.45	.51
	feel like getting behind in work	.10	.32	.20	.42	.15	.37
	feel like doing less pre-planning	.20	.42	.30	.48	.25	.44
	work slower as usual	.70	.82	.50	.71	.60	.75
	not knowing a/c on frequency	.10	.32	.20	.42	.15	.37
	forgetting routine co-ordination	.00	.00	.00	.00	.00	.00
	react slower as usual	.40	.70	.30	.48	.35	.59
	pay less attention to detail	.60	.84	.50	.53	.55	.69
	surprised by call	.20	.42	.50	.53	.35	.49
	miss a call	.20	.42	.40	.52	.30	.47
	feel easily distracted	.50	.53	.78	.83	.63	.68
	look for traffic that calls in	.60	.70	.70	.67	.65	.67
	overlooking obvious problems	.00	.00	.10	.32	.05	.22
	do less scanning	1.10	1.37	1.10	.99	1.10	1.17
TOT	not feel focused anymore	.60	.75	.55	.83	.58	.78
	feel complacent	1.79	1.93	2.68	2.03	2.24	2.01
	ask to repeat call from a/c	.65	.88	.85	.75	.75	.81
	not understand R/T	.35	.67	.37	.50	.36	.58
	spot a conflict only 1-6 minutes before	.65	1.46	.55	1.00	.60	1.24
	feel like working less precise	.65	.88	.79	1.40	.72	1.15
	make small mistakes (e.g. input errors)	.55	.69	.55	.51	.55	.60
	feel like getting behind in work	.10	.31	.20	.41	.15	.36
	feel like doing less pre-planning	.40	.60	.42	.61	.41	.59
	work slower as usual	.45	.69	.50	.76	.48	.72
	not knowing a/c on frequency	.20	.41	.25	.44	.23	.42
	forgetting routine co-ordination	.00	.00	.05	.22	.03	.16
	react slower as usual	.50	.61	.35	.59	.43	.59
	pay less attention to detail	.55	.76	.50	.51	.53	.64
	surprised by call	.20	.41	.45	.51	.33	.47
	miss a call	.25	.55	.30	.47	.28	.51
	feel easily distracted	.45	.51	.56	.70	.50	.60
	look for traffic that calls in	.70	.73	.75	.72	.73	.72
	overlooking obvious problems	.15	.37	.15	.37	.15	.36
	do less scanning	.75	1.07	.80	.83	.78	.95

Table C-7: Statistical analysis for further performance ratings

a) ask to repeat call from a/c

Source	Numerator df	Denominator df	F	p
Density	1	25	.81	.377
Repetitiveness	1	25	.12	.738
Density x Repetitiveness	1	25	2.644	.116

b) not understand R/T

Source	Numerator df	Denominator df	F	p
Density	1	25	.00	.989
Repetitiveness	1	25	.00	.989
Density x Repetitiveness	1	25	3.12	.090

c) spot a conflict only 1-6 minutes before

Source	Numerator df	Denominator df	F	p
Density	1	25	.07	.787
Repetitiveness	1	25	.14	.685
Density x Repetitiveness	1	25	.45	.510

d) feel like working less precise

Source	Numerator df	Denominator df	F	p
Density	1	25	.02	.898
Repetitiveness	1	25	.49	.493
Density x Repetitiveness	1	25	2.26	.146

e) make small mistakes (e.g. input errors)

Source	Numerator df	Denominator df	F	p
Density	1	24	.05	.823
Repetitiveness	1	25	1.74	.199
Density x Repetitiveness	1	25	.56	.460

f) feel like getting behind in work .

Source	Numerator df	Denominator df	F	p
Density	1	26	.38	.545
Repetitiveness	1	26	.00	.967
Density x Repetitiveness	1	26	.00	.967

g) feel like doing less pre-planning

Source	Numerator df	Denominator df	F	p
Density	1	26	.00	.968
Repetitiveness	1	26	2.54	.123
Density x Repetitiveness	1	26	.14	.713

h) work slower as usual

Source	Numerator df	Denominator df	F	p
Density	1	24	.03	.866
Repetitiveness	1	24	.72	.404
Density x Repetitiveness	1	24	.72	.404

i) not knowing a/c on frequency

Source	Numerator df	Denominator df	F	p
Density	1	24	.00	.958
Repetitiveness	1	25	2.51	.126
Density x Repetitiveness	1	25	.63	.437

j) forgetting routine co-ordination

Source	Numerator df	Denominator df	F	p
Density	1	26	.84	.381
Repetitiveness	1	26	.84	. 381
Density x Repetitiveness	1	26	.84	. 381

k) pay less attention to detail

Source	Numerator df	Denominator df	F	p
Density	1	24	1.05	.315
Repetitiveness	1	24	.01	.922
Density x Repetitiveness	1	24	.01	.922

l) surprised by call

Source	Numerator df	Denominator df	F	p
Density	1	24	2.93	.100
Repetitiveness	1	24	.01	.938
Density x Repetitiveness	1	24	.50	.486

m. miss a call

Source	Numerator df	Denominator df	F	p
Density	1	25	.02	.886
Repetitiveness	1	25	.02	.887
Density x Repetitiveness	1	25	1.61	.216

n. look for traffic that calls in

Source	Numerator df	Denominator df	F	p
Density	1	24	.05	.824
Repetitiveness	1	25	.18	.678
Density x Repetitiveness	1	25	.00	.992

o. overlooking obvious problems

Source	Numerator df	Denominator df	F	p
Density	1	25	.05	.824
Repetitiveness	1	25	2.51	.125
Density x Repetitiveness	1	25	.48	.495

Table C-8: Descriptive statistics for traffic-related indicators

			Traffic Load				Group Total	
			low		high			
			M	SD	M	SD	M	SD
Repetitiveness	non repetitive	traffic repetitiveness	3.73	1.36	2.97	1.00	3.35	1.23
		traffic difficulty	4.27	1.32	4.50	.91	4.38	1.11
		traffic density	4.23	1.27	4.27	.93	4.25	1.08
		traffic complexity	4.37	1.39	4.60	.83	4.48	1.12
		traffic routine	4.17	1.35	3.30	1.31	3.73	1.37
	repetitive	traffic repetitiveness	4.63	1.85	3.40	1.36	4.02	1.70
		traffic difficulty	2.10	1.10	2.73	1.29	2.42	1.21
		traffic density	2.52	1.13	3.33	.96	2.93	1.10
		traffic complexity	2.30	1.10	2.97	1.25	2.63	1.20
		traffic routine	5.15	1.49	4.13	1.25	4.64	1.44
Group Total	traffic repetitiveness		4.18	1.65	3.18	1.18	3.68	1.50
	traffic difficulty		3.18	1.62	3.62	1.42	3.40	1.52
	traffic density		3.38	1.46	3.80	1.03	3.59	1.27
	traffic complexity		3.33	1.62	3.78	1.33	3.56	1.48
	traffic routine		4.66	1.48	3.72	1.32	4.19	1.46

Table C-9: Statistics for additional variables

	N	M	SD
ACQ: HOM	9	8.11	2.31
ACG:HOP	9	9.44	2.35
ACG: HOT	9	10.67	.86
MES	10	53.19	4.76
Satisfaction	10	2.86	.49

Average values and ranges in the subscales for the assessment of work satisfaction ranging from 1 (very satisfied) to 5 (not at all satisfied).

	Minimum	Maximum	M	SD
Satisfaction with Information and communication	2.33	4.67	3.43	.83
Satisfaction with how demanding your job is	1.00	3.50	2.02	.69
Satisfaction with the contact to your colleagues	1.25	3.75	2.05	.79
Satisfaction with the relationship to your nearest boss	2.00	4.00	2.80	.76
Satisfaction with organization and management	2.00	5.00	3.63	.89
Satisfaction with the working conditions	1.67	4.67	3.43	.96
Satisfaction with my freedom in decision-making	1.00	3.67	2.36	.92
Satisfaction with the payment I receive	3.67	5.00	4.03	.48
Satisfaction with working- and vacation times	1.80	3.80	2.91	.68
Satisfaction with the general working conditions	1.67	4.00	2.83	.81

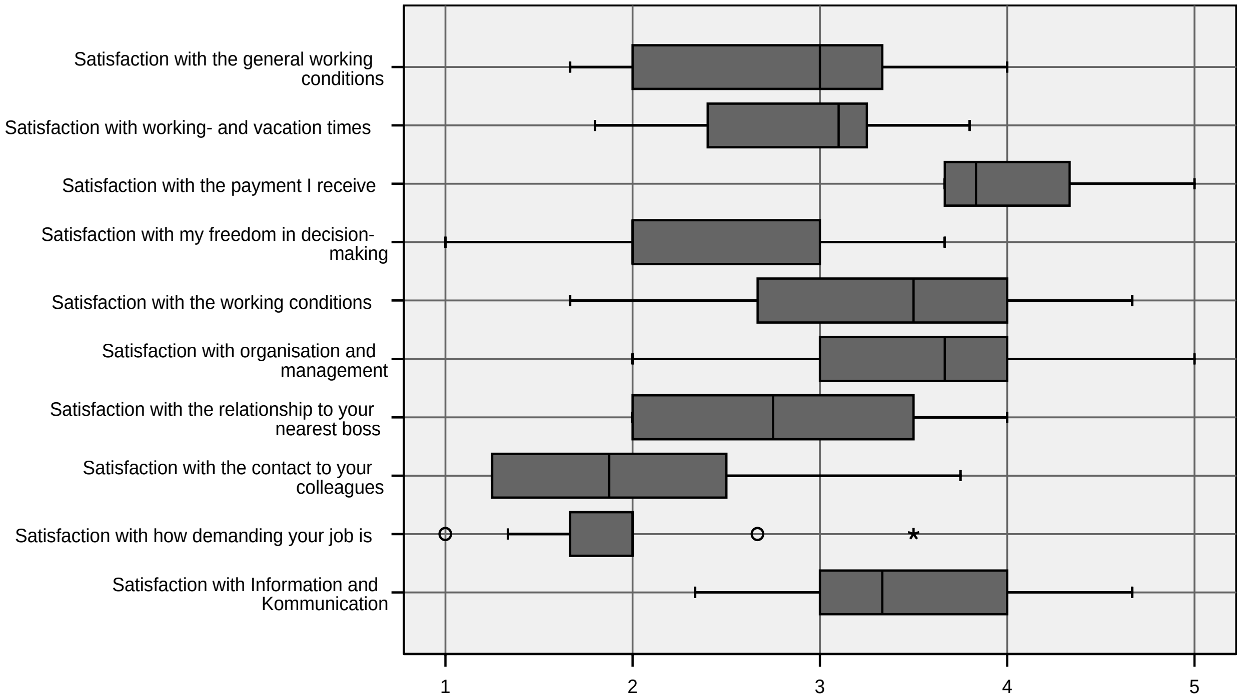


Table C-10: Descriptive statistics traffic load groups in subjective and traffic indicators

			Traffic Load					
			low		medium		high	
			M	SD	M	SD	M	SD
Section	1	attentiveness	3.07	1.49	3.43	.85	4.45	1.51
		concentration	3.36	1.65	3.29	.91	4.42	1.83
		effort	2.86	1.56	2.57	1.22	4.33	1.83
		workload	2.60	.83	2.87	.78	3.60	.87
		feeling of monotony	3.71	2.23	4.29	1.27	2.92	1.98
		boredom	3.71	2.30	3.71	1.77	2.64	1.69
	2	attentiveness	3.07	1.03	3.53	1.51	5.11	.78
		concentration	3.07	1.28	3.80	1.70	5.30	.82
		effort	2.53	1.36	3.13	1.92	5.00	1.05
		workload	2.91	.92	2.89	.96	3.72	.45
		feeling of monotony	4.07	1.79	3.67	2.38	2.20	1.32
		boredom	4.13	1.96	3.53	2.23	2.00	1.32
	3	attentiveness	3.80	1.32	3.36	1.69	4.41	1.50
		concentration	3.70	1.70	3.55	1.86	4.67	1.50
		effort	3.30	1.77	2.91	1.81	4.50	1.47
		workload	3.14	.90	2.80	1.15	3.69	.63
		feeling of monotony	3.70	2.11	3.73	2.24	2.67	1.68
		boredom	3.60	2.12	3.64	2.16	2.53	1.70

			Traffic Load					
			low		medium		high	
			M	SD	M	SD	M	SD
Section	1	traffic density	3.00	1.41	3.07	1.07	4.33	1.44
		traffic complexity	3.00	1.57	2.71	1.27	4.33	1.72
		traffic repetitiveness	4.29	1.82	3.36	1.34	3.50	1.38
		traffic routine	4.29	1.77	4.43	1.22	4.00	1.35
		traffic difficulty	2.86	1.41	2.64	1.08	4.25	1.82
	2	traffic density	2.73	1.10	3.33	1.50	4.90	1.20
		traffic complexity	2.73	1.22	3.07	1.62	5.20	1.14
		traffic repetitiveness	3.67	1.63	3.87	1.88	3.40	1.17
		traffic routine	4.87	1.30	4.33	1.76	3.30	1.49
		traffic difficulty	2.67	1.45	3.00	1.77	4.90	.99
	3	traffic density	3.50	1.72	3.09	1.76	4.56	1.29
		traffic complexity	3.60	1.65	3.09	2.02	4.67	1.33
		traffic repetitiveness	3.10	1.66	4.64	2.01	3.28	1.36
		traffic routine	4.00	1.76	4.55	1.51	3.61	1.42
		traffic difficulty	3.20	1.99	3.09	2.02	4.33	1.28

Table C-11: Descriptive statistics of critical states (SOF) and mood subscales (UWIST) during the shift (Descriptive Hypothesis 5)

			M	SD
Monotony	Day 1	Morning	1.99	.29
		After WP 1	2.03	.24
		After WP 2	2.06	.39
	Day 2	Morning	2.04	.28
		After WP 1	1.95	.25
		After WP 2	2.14	.29
			M	SD
Fatigue	Day 1	Morning	1.99	.29
		After WP 1	2.03	.24
		After WP 2	2.06	.39
	Day 2	Morning	2.04	.28
		After WP 1	1.95	.25
		After WP 2	2.14	.29
			M	SD
Satiation	Day 1	Morning	1.38	.24
		After WP 1	1.58	.55
		After WP 2	1.51	.45
	Day 2	Morning	1.59	.35
		After WP 1	1.63	.44
		After WP 2	1.84	.43
			M	SD
Stress	Day 1	Morning	1.48	.36
		After WP 1	1.53	.24
		After WP 2	1.57	.50
	Day 2	Morning	1.48	.34
		After WP 1	1.57	.40
		After WP 2	1.62	.37

			M	SD
Hedonic tone	Day 1	Morning	2.58	.23
		After WP 1	2.57	.22
		After WP 2	2.49	.17
	Day 2	Morning	2.68	.26
		After WP 1	2.60	.23
		After WP 2	2.61	.29
Tense arousal	Day 1	Morning	2.93	.21
		After WP 1	2.82	.25
		After WP 2	2.79	.19
	Day 2	Morning	2.85	.19
		After WP 1	2.79	.14
		After WP 2	2.87	.18
Energetic arousal	Day 1	Morning	2.16	.32
		After WP 1	2.08	.41
		After WP 2	2.24	.33
	Day 2	Morning	2.01	.31
		After WP 1	2.03	.34
		After WP 2	1.90	.26

Table C-12: Descriptive statistics changes in traffic density (Descriptive Hypothesis 6)

			M	SD
Shift H-L	no	motivation	4.17	1.53
		Monotony indicator (z-value)	-.01	.84
	yes	motivation	3.83	1.55
		Monotony indicator (z-value)	.00	.80
Shift L-H	no	motivation	3.98	1.52
		Monotony indicator (z-value)	.15	.80
	yes	motivation	4.18	1.58
		Monotony indicator (z-value)	-.25	.80

Table C-13: Personally relevant situations

Code	Day	WP	Description of situation	Traffic situation	Additional factors	Type of task executed	Deviation from routine	SET-Rating S	E	T	W
14	1	1	Working in the lower sector (up to 295) when an aircraft called in at FL 330. This was outside the own filter range. with the "Quick look" button the aircraft became visible	moderate density	team factors: other country's mistake	information collection; request	routine	2	1	-1	2
24	2	2	Aircraft which was previously not in my sector (FL 340) suddenly climbed to FL 350 which is my sector already. The separation tool helps quickly to solve such situations It helps vectoring very well.	moderate density	team factors	information collection; request; radar monitoring	routine	2	1	0	4
25	2	2	Simple military over flight, not part of the NATO, so they need a special clearance to fly over Austria, and I guess it was a misunderstanding, but the supervisor from Vienna called our supervisor and said that this flight had no entry permission, but it was a bit late. At the time we got the call from Vienna, it was 2 min before entering the Austrian airspace. I was at the end of this line, so when I got the message, there was one minute to go. So the procedure was that I had to tell the pilot to hold at the exit point to Austria, and he has to wait for entry permission. And at the very last time I gave him the instruction to report over Austria, but at the time he passed the point, so it is a very difficult procedure if he passed the point to enter the holding, there is a procedure to hold to that, so it was senseless instruction, in addition to this, the pilot informed me that he has entry permission. So to discuss this problem with Vienna, I sent him to Vienna frequency, so he partly entered to holding without making the circle, and at mid-time of this holding procedure he went away to the	low density	communication factors; team factors	radar monitoring	strong deviation	-1	0	-2	1

Code	Day	WP	Description of situation	Traffic situation	Additional factors	Type of task executed	Deviation from routine	SET-Rating			
								S	E	T	W
			normal route, so they found the permission in Vienna.								
16	1	1	And there is a special event, there were two AC, was an AFR sent by Vienna with destination Budapest at the same position at 310 and 330 and the aircraft was asked to get from Vienna at high speed. But when I got them they were at the same speed. So that's why the conception had to change because of this situation. but nothing special.	high density	team factors: in Vienna there was a situation that has changed. that is why we had the changes	coordination; because I had to coordinate with TOP sector who was waiting for the traffic to climb to their level request; Comm with planner and adjacent sector; electronic strip work; radar monitoring	routine				
17	1	2	sequencing problem with arriving traffic in Budapest	complex traffic		Coordination (VH Appraoch); Information collection IAS; request to direct. any direct; communication with planner and adjacent sector; electronic strip work; radar monitoring;	routine	2	1	0	3

Appendix D: ADDITIONAL INFORMATION

D.1. INTRODUCTION TO SCIENTIFIC EXPERIMENTS AND HYPOTHESIS TESTING

The general goal of a scientific experiment is to test a hypothesis which has been generated based on previous observations or theory. A hypothesis is a specific statement that predicts a result and can be verified or falsified (*cf.* Popper, 1935, reprinted 2002, for a discussion of principles of scientific discovery). Through the manipulation of independent variables, an experimenter tries to determine the effect of a condition, the independent variable, on the individual's behaviors, the dependent variables, through the control of all relevant factors. Generally stated, research hypotheses are transferred to the statistical hypothesis, which includes a statistical measure and can thus be submitted to a statistical test. It is however noted that a research hypothesis corresponds to the alternative hypothesis and that a researcher generally tries to disprove a null hypothesis.

A statistical hypothesis is set up in form of a null hypothesis, which predicts that an observed difference between conditions in a defined population is only due to chance and not to a systematic effect of the independent variable. In other words, no significant difference between observed conditions is found. A predicted difference in two or more conditions is formulated in the alternative hypothesis and assessed with calculating the probability (p-value) that the difference between two populations is due to random. If the probability that an observed effect is due to chance is lower than the defined margin called the alpha level. it is referred to as *statistically significant*. Depending on the size of the p-value in relation to the conventionally defined alpha level, a null hypothesis is retained or can be rejected in support for presuming the alternative hypothesis.

An additional aspect is related to the risks of erroneously accepting or rejecting a wrong hypothesis. The type I error called alpha represents the risk of rejecting the null hypothesis if it is true and is conventionally set to p(alpha) = 0.05. The risk to not reject the null hypothesis even if it is false and thus not to detect a significant true effect in the population is called type II or beta error, conventionally set to p(beta) = 0.20. A way to determine if the experimental design is efficient to detect a significant true effect is power analysis. This occurs through the consideration of the effect size, which is the size of the statistically significant difference between conditions, the sample size and alpha and beta errors. Ideally, the power of an experiment should be $p = 0.80$.

D.2. AVAILABILITY OF FURTHER MATERIALS

The following additional materials are available in an electronic data format:

- Materials and Instructions for all studies
- Raw data files
- Statistical analysis procedures

Please contact the following address:

Karl-Franzens-Universität Graz
Institut für Arbeits-, Organisations- und Umweltpsychologie
Universitätplatz 2/III
8010 Graz
AUSTRIA